THE IMPACT OF THE TERMINATION OVERRIDE MUTATION

ON THE ACTIVITY OF SSU72

Neil Andrew McCracken

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Biochemistry and Molecular Biology,
Indiana University

December 2016

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis
Committee

Amber L Mosley, PhD. Chair

Mark G. Goebl, PhD.

Ronald C. Wek, PhD.

ACKNOWLEDGEMENTS

The following people played an instrumental role in the completion of this work and I am so very thankful for each of them.

- Dr. Amber Mosley: It was such a pleasure to be part of your lab family over the past two years. Your constant encouragement coupled with critique has allowed me to grow as a scientist more than any other period of time in my life. I also consider you a friend and will always appreciate your value of my opinion and how our discussions were like those of colleagues.

- Dr. Mark Goebl: Your roles as member of my committee and as a lecturer in my classes has taught me the importance of simplicity and practicality. I am also thankful for your help and guidance two years ago when I was first considered the graduate program.

- Dr. Ron Wek: You are my most favorite lecturer in all of my years of undergrad, graduate school and 15 years in industry. You make understanding the topic of protein expression achievable to a chemical engineer. At the same time, our discussions made me realize how much there really is for me to learn.

- Whitney Smith-Kinnaman: I feel like you were always looking out for me in the lab and the lab would not be a family without you.

- Dr. Jerry Hunter: I will never forget our weekends in the lab together when I first started the Master's program. Your willingness to share lab methods with me were essential to this work.

- My lab mates Sarah Peck, Jose Victorino, Dr. Mel Fox, Asha Boyd and Gabi Mazur: You made school fun.

- My daughter Ruby: You bring me so much joy and I am so very lucky to have you call me daddy. I am thankful for each and every day that I have known you. While I never want you to grow up, I look forward to the days ahead as you become the absolutely amazing woman you were made to be.

- My son Nolan: You will always be my boy. I will never forget the day you were born and how it felt to hold you in my arms. You make me proud and I will always be your number one fan. Your laugh and sense of humor are contagious and your caring for other people will surely carry you far in life.

- And most importantly, my wife Becca: You will always be my love and best friend. The late nights and weekends in the lab or studying at the coffee shop would not have happened without you. You are the strongest woman I know. It is your strength and selflessness that have brought us through challenges while it is your joy and spirit that have made our walk through life enjoyable.

Neil Andrew McCracken

THE IMPACT OF THE TERMINATION OVERRIDE MUTATION ON THE ACTIVITY OF SSU72

Ssu72, an RNA Pol II CTD phosphatase that is conserved across eukaryotes, has been reported to have a wide array of genetic and physical associations with transcription factors and complexes in RNA transcription. Catalytic mutants of Ssu72 are lethal across many eukaryotes, and mutations to non-catalytic sites in SSU72 phosphatase have been shown to lower function. One spontaneous mutation of the SSU72 gene in *Saccharomyces cerevisiae* (A to C nucleotide mutation resulting in an L84F mutation in the coded protein) was shown to have transcription termination deficiency (termination override or TOV). This SSU72 mutation was suggested by Loya et al. to cause a lowering of the phosphatase activity of the protein and consequently affect proper termination. In research reported herein, an investigation was completed through *in-vitro* and *ex-vivo* approaches with the goal of understanding the impact of the SSU72 TOV mutation on the observed phenotype in *S. cerevisiae*. It can be concluded from work presented in this report that the SSU72 TOV mutation does not cause a decrease in *in-vitro* phosphatase activity as compared to wild type. Evidence presented even suggests an increase in phosphatase activity as compared to wild type Ssu72. One model for the observed responses in transcription termination is that the phenylalanine substitution in Ssu72 leads to cooperative interactions with proline residues in the CTD. It is proposed that the corresponding increase in Ssu72 phosphatase activity limits RNA Pol II CTD association with termination factors, such as Nrd1, thus causing deficient transcription termination.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3C: Chromosome Conformation Capture

A: Adenine

C: Cytosine

ChIP: Chromatin Immunoprecipitation

CPF: Cleavage and Polyadenylation Factor

CTD: C-terminal Domain

DiFMUP: 6,8-difluoro-4-methylumbelliferyl phosphate

DNA: Deoxyribonucleic Acid

dsDNA: Double Stranded Deoxyribonucleic Acid

FLAG: Affinity tag with DYKDDDDK sequence

GFP: Green Fluorescent Protein

GST: Glutathione S-Transferase

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IT: Intergenic Terminator

L84F: Leucine to Phenylalanine Mutation at Residue 84 of Ssu72

LB: Lysogeny Broth

LC: Liquid Chromatography

miRNA: Micro Ribonucleic Acid

mRNA: Messenger Ribonucleic Acid

MS: Mass Spectrometry

MudPIT: Multi-Dimensional Protein Identification Technology

MWCO: Molecular Weight Cut-off

ncRNA: Non-coding Ribonucleic Acid

NTP: Nucleotides

PAF: a complex which includes Paf1, Ctr9, Cdc73, Trf1 and Leo1

PNPP: Para-Nitropheynyl Phosphate

Pol: RNA Polymerase

rcf: Relative Centrifugal Force

pRb: retinolblastoma tumor suppressor

RNA: Ribonucleic Acid

rRNA: Ribosomal Ribonucleic Acid

RT-qPCR: Real Time Quantitative Reverse Transcription Polymerase Chain Reaction

SDS-PAGE: Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis

snRNA: Small Nuclear Ribonucleic Acid

snoRNA: Small Nucleolar Ribonucleic Acid

S.O.C.: Super Optimal Broth with Catabolite Repression

Tm: Melting Temperature

TOV: Termination Override

tRNA: Transfer Ribonucleic Acid

WT: Wild Type

INTRODUCTION

Transcription of DNA to RNA is the prerequisite step for all cellular functions as it creates the message (mRNA) that is later translated into proteins. In many cases, transcription produces non-coding RNAs (i.e. miRNA, lncRNA, snRNA and snoRNA) that are used for either the execution of transcriptional steps or other regulatory processes. Transcription of RNAs in eukaryotes is executed by three polymerases; RNA Polymerase I (hence Pol I), RNA Polymerase II (hence Pol II) and RNA Polymerase III (hence Pol III) (Cramer et al., 2008). Pol I is responsible for transcribing ribosomal RNA (excluding 5S rRNA), whereas Pol III transcribes transfer RNA (tRNA), 5S ribosomal RNA, and some small RNAs. Pol II is responsible for generating not only RNAs that can code for protein, pre-messenger RNA (pre-mRNA), but also other non-coding RNAs that include small nuclear RNA (snRNA) and microRNA (miRNA).

The transcription process by Pol II has been well characterized despite its complexity. DNA transcription by Pol II is traditionally broken down into the three phases of initiation, elongation and termination. Initiation corresponds with the docking of Pol II to the upstream portion of the gene that is to be transcribed. The genetic initiation region contains a variety of elements such as promotors, enhancers and silencers that stimulate or repress start site recognition by Pol II (Breathnach & Chambon, 1981; O'Shea-Greenfield & Smale, 1992; Riethoven, 2010; Smale, Schmidt, Berk, & Baltimore, 1990). Other proteins, known as transcription factors, also act in partnership with Pol II to facilitate DNA site recognition, DNA binding promotion or interaction stabilization (Roeder, 1991). The general factors that are associated with initiation are TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. After Pol II has docked with the DNA template and the pre-initiation complex (PIC) is formed, the hydrogen bonds between nucleotides are

broken by the helicase activity of TFIIH (T. K. Kim, Ebright, & Reinberg, 2000). Melting of nucleotide bonds results in the formation of an open complex (Wang, Carey, & Gralla, 1992), which allows RNA Pol II to initiate synthesis of the nascent RNA. Factors such as TFIIS, SWI/SNF, Spt4, Spt6, and FACT (Sims, Belotserkovskaya, & Reinberg, 2004) are then recruited to Pol II in order to progress to the next phase of transcription elongation (Bentley, 2002), while other factors are thought to remain behind but near the promoter to act as guides for re-initiation (Yudkovsky, Ranish, & Hahn, 2000; Zawel, Kumar, & Reinberg, 1995).

The transition from initiation to the elongation phase has been suggested to be regulatory step in gene expression (T. H. Kim et al., 2005). At the point of transition, Pol II can go through steps of promoter escape, pausing, and “productive elongation” (Saunders, Core, & Lis, 2006). The success of the initiation to elongation transition is regulated by the presence of transcription activators/factors such as SWI/SNF (Brown, Imbalzano, & Kingston, 1996). If Pol II is indeed able to escape the promoter region and continue to elongation, nucleotides (NTPs) are added to the nascent transcript. In the process of elongation, Pol II is accompanied by factors that are both passive and active. Passive factors (e.g. the PAF complex) associate with Pol II for the duration of elongation while active factors (e.g. TFIIF) associate for short periods of time during elongation (Mayer et al., 2010; Sims et al., 2004).

Prevention of downstream (and potentially undesired) gene transcription during transcript synthesis is of utmost importance during elongation. Characterization of the transition from transcript elongation to termination is complicated by the fact that the

proposed model depends on whether the transcript is coding or non-coding (Porrua & Libri, 2015). In the case of mRNA, coordinated with Pol II separation from the DNA template, a chain of adenines is added to the 3'-end of the transcript (Richard & Manley, 2009). Non-coding transcripts on the other hand, can require the participation of factors such as Nrd1, Nab3 and Sen1 (NNS) to aid in the termination process (Vasiljeva, Kim, Mutschler, Buratowski, & Meinhart, 2008). In general, the process of transcription termination requires the association of transcription factors with Pol II to regulate and aid in the disengagement of the polymerase from the DNA. A general pictorial description of the transcription process is shown in Figure 1.

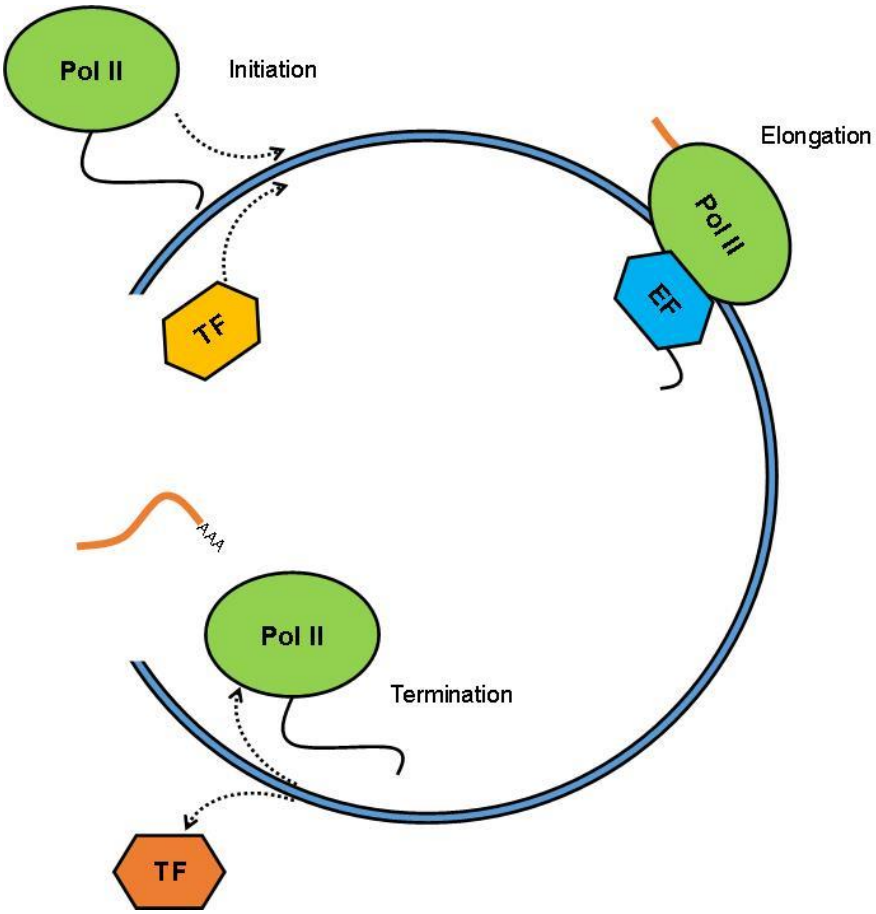


Figure 1. Model of Pol II transcription phases (yellow TF = Transcription Factors, blue EF = Elongation Factors, red TF = Termination Factors)

Many co-transcriptional activities are orchestrated by the C-terminal Domain (CTD) of subunit 1 of the 12 unit Pol II (S. Buratowski, 2009). The CTD is composed of heptad repeats (YSPTSPS) numbering 26 in budding yeast and 52 in vertebrates where the length of this unstructured domain has been associated with the complexity of the organism (Corden, Cadena, Ahearn, & Dahmus, 1985; Hsin & Manley, 2012). Besides initiation, elongation and termination, the CTD has been associated with other features of gene expression, including polyadenylation, 5'-capping, histone methylation, and pre-mRNA splicing (Stephen Buratowski, 2003; McCracken, Fong, Yankulov, et al., 1997). In order for the polymerase to participate in this wide array of tasks, the CTD is phosphorylated at serine and tyrosine sites at different times during transcription to accommodate docking of proteins that assist in the many functions of gene expression that were discussed above. Chromatin Immunoprecipitation (ChIP) studies have revealed that serine 5 is predominantly phosphorylated at the beginning of gene transcription, whereas serine 2 and serine 7 of the CTD is phosphorylated toward when transcription elongation proceeds towards the 3'-end of the gene (Chapman et al., 2007; P. Komarnitsky, E.-J. Cho, & S. Buratowski, 2000; Mayer et al., 2010; Phatnani & Greenleaf, 2006).

An illustration of the varying states of phosphorylation that are present through the phases of ranging from initiation (Jeronimo, Bataille, & Robert, 2013) to termination is shown in Figure 2 (where the deeper color indicates a phosphorylated serine).

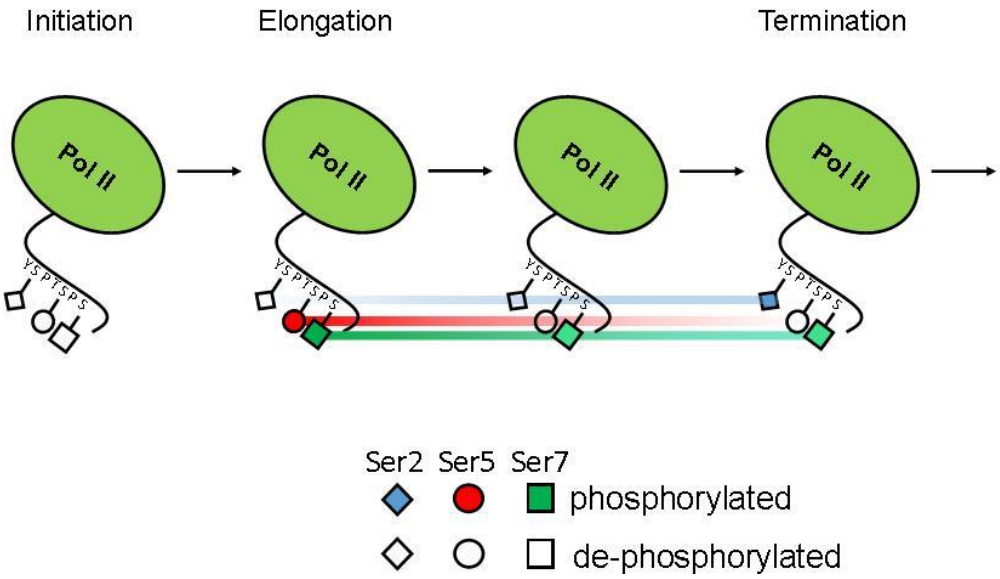


Figure 2. Varying states of CTD phosphorylation based on stage of transcription

One protein kinase that engages with Pol II, TFIIH, has specificity for phosphorylation of serine 5 (Philip Komarnitsky et al., 2000), and serine 5 has been shown through multiple methods to associate with factors participating the capping of pre-mRNAs (P. Komarnitsky, E. J. Cho, & S. Buratowski, 2000; McCracken, Fong, Rosonina, et al., 1997). On the other hand, the protein kinase CTDK-I has been reported to engage with factors associating with the 3' end of the mRNA and a subunit of a cleavage factor (CF IA) that binds to phosphorylated serine 2 in the CTD (Licatalosi et al., 2002). Furthermore, other subunits of the CF1A cleavage factor, Rna15, Rna14, Pcf11p, were also shown to associate with phosphorylated CTD (Barilla, Lee, & Proudfoot, 2001; B. Dichtl et al., 2002; Kyburz, Sadowski, Dichtl, & Keller, 2003).

Just as phosphorylation of the serine residues of the CTD affects transcription, it is also understood that de-phosphorylation is necessary to maintain proper function of RNA Pol II. CTD phosphatases are responsible for removal of the phosphate at serine and threonine residues and some that have been identified include Fcp1 (Archambault et al.,

1997), Scp1 (Yeo, Lin, Dahmus, & Gill, 2003) and Rtr1 (Mosley et al., 2009). Another CTD phosphatase, SSU72, was initially described as having genetic association with transcription factor IIB (TFIIB) in *S. cerevisiae* (Sun & Hampsey, 1996). Specifically, it was shown in this initial report by Sun et al. that a mutant of SSU72 was found to be an enhancer of a SUA7 defect[a] and that cysteine residues and the N-terminus of Ssu72 were essential for function. Physical interactions between Ssu72 and RNA Pol II were also reported based on co-immunoprecipitation assays (Pappas & Hampsey, 2000). Association of Ssu72 with the 3'- end of DNA has been found as the protein was found to be stably associated with the RNA cleavage and polyadenylation factor (CPF) and was described to be a possible "bridge" between CPF subunits (Bernhard Dichtl et al., 2002; Eduard Nedea et al., 2003). Overall, Ssu72 has the unique distinction of being associated with not only downstream transcription elongation but also the transition between elongation and termination. The mixed functionality of Ssu72 was further supported when the protein was described to have a phosphatase role independent of its CPF association. The identification of the phosphatase activity was based not only on structural motif of the protein (CX5R) (Meinhart, Silberzahn, & Cramer, 2003) but also on its specificity of Ssu72 for serine 5 of the CTD heptad repeats (Krishnamurthy, He, Reyes-Reyes, Moore, & Hampsey, 2004) and later specificity for serine 7 (Bataille et al., 2012; Xiang, Manley, & Tong, 2012; D. W. Zhang et al., 2012) . Furthermore, it was reported that Ssu72 preferentially binds to the CTD of Pol II when cis-proline is present in the heptad repeats (Mayfield et al., 2015b; Werner-Allen et al., 2011) as opposed to trans-proline. The Ssu72 protein was more recently studied in Caenorhabditis elegans where it was found to not be requisite for viability (F. Chen et al., 2015). At the same

[a] *SUA7* encodes for the transcription factor, TFIIB. The Ssu72 protein was found with other suppressors of a *SUA7* defect and was therefore designated as a suppressor of *SUA7*. The *SUA7* designation is a misnomer due to the fact that it was actually shown to be associated with enhancing a *SUA7* defect.

time, Ssu72 was shown in *C. elegans* to be connected with regulation of alternative polyadenylation sites of pre-RNAs during neuronal development (F. Chen et al., 2015).

Many of the functions described so far are conserved in mammals (St-Pierre et al., 2005) and have consequently led to investigations that are more directed to translational applications. For instance, it was reported that one of the transcription elongation factors for the HIV-1 virus (Tat) associates directly with Ssu72 thus linking the phosphatase with viral gene expression in host cells (Y. Chen et al., 2014). Another report (St-Pierre et al., 2005) used a yeast two-hybrid approach to show that Ssu72 physically associates with the human retinoblastoma tumor suppressor (pRb); a protein whose inactivation in human cells contributes to different forms of cancer (Liu, Dibling, Spike, Dirlam, & Macleod, 2004). Alignment of the Ssu72 protein in yeast with the human orthologue is shown below in Figure 3 and illustrates the conservation among these phosphatases as exemplified by the catalytic cysteine residue (highlighted in yellow) along with the other consensus in the sequences.

```
Q9NP77 SSU72_HUMAN 1    ---MPSSPLRVAVVCSSNQNRSMEAHNILSKRGFSVRSFGTGTHVKLPGPAPDKPNVYDF   57
P53538 SSU72_YEAST 1    MPSHRNSNLKFCTVCASNNNRSMESHKVLQEAGYNVSSYGTGSAVRLPGLSIDKPNVYSF   60
                             .* *:...**:**:*****:*::*.: *:.* *:***: *:*** : ******.*

Q9NP77 SSU72_HUMAN 58   KTTYDQMYNDLLRKDKELYTQNGILHMLDRNKRIKPRPERFQNCKDLFDLILTCEERVYD  117
P53538 SSU72_YEAST 61   GTPYNDIYNDLLSQSADRYKSNGLLQMLDRNRRLKKAPEKWQESTKVFDFVFTCEERCFD  120
                        * *:::***** :. : *..**:*:*****:*:*  **::*:...:**:::***** :*

Q9NP77 SSU72_HUMAN 118  QVVEDLNSREQETCQPVHVVNVDIQDNHEEATLGAFLICELCQCIQH--------TEDME  169
P53538 SSU72_YEAST 121  AVCEDLMNRGGKLNKIVHVINVDIKDDDENAKIGSKAILELADMLNDKIEQCEKDDIPFE  180
                        * *** .*  :  : ***:****:*:.*:*.:*:  * **.: ::.          :*

Q9NP77 SSU72_HUMAN 170  NEIDELLQEFEEKSGR-TFLHTVCFY  194
P53538 SSU72_YEAST 181  DCIMDILTEWQSSHSQLPSLYAPSYY  206
                        : * ::* *::.. .:   *:: .:*
```

Figure 3. Alignment between Ssu72 protein sequence in Baker's yeast and *Homo sapiens*

Association of Ssu72 with both the 5'- and 3'- ends of transcribed genes has contributed to research on the role of Ssu72 in gene looping. It has been shown by several methods, including: genetic interactions between Ssu72 and the initiation factor TFIIB (Sun & Hampsey, 1996), electron microscopy (Morgan, 2002) and chromosome conformation capture (3C) (Tan-Wong et al., 2012), that the promoter regions and associated initiation factors can associate with the 3' ends of the gene and its corresponding factors (such as the polyadenylation complex). The association between the two distal regions of a gene not only can aid in the efficiency of the transcriptional process for highly expressed genes but also potentially act to regulate the transcription of the sense versus antisense DNA template strand. Specifically, it was observed through 3C analysis that mutation of Ssu72 prevents formation of gene loops across a specific gene and also results in an increase in ncRNA expression (Tan-Wong et al., 2012). Therefore, Ssu72 was implicated in regulating the balance between transcription of coding and non-coding transcripts. An adapted model for Pol II transcription featuring Ssu72 is shown in Figure 4.

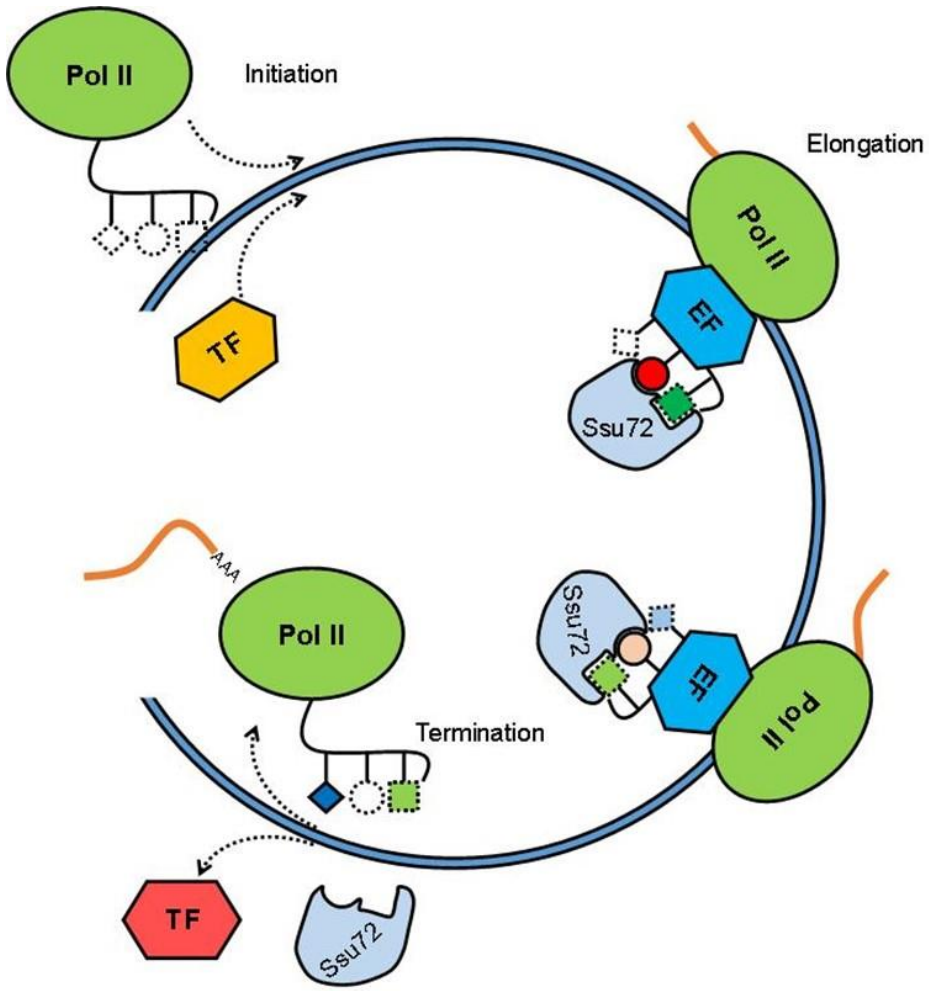


Figure 4. Model of Pol II transcription phases with CPF and Ssu72

As with many proteins, our understanding Ssu72 relies on observing the impact of conditional and/or non-lethal mutations of SSU72 on transcription. One such mutation (Loya, O'Rourke, & Reines, 2012) was found in a genetic screening of *S. cerevisiae* that had been transformed with a unique expression construct. Specifically, the introduced vector contained a strong intergenic terminator (IT) flanked on either side by a galactose inducible promoter and a GFP reporter. The goal of the experiment was to isolate mutants with ineffective termination as indicated by a higher level of fluorescence in FACS screening. Termination Override (TOV) mutants were sorted and two

spontaneous mutants were identified that coincidentally each contained identical mutations in the SSU72 gene. Specifically, an A to C mutation in the nucleotide sequence led to expression of the Ssu72 phosphatase with leucine at residue 84 changed to a phenylalanine. Association between the L84F mutant Ssu72 protein and the corresponding termination deficient phenotype *in-vivo* was unclear but it was suggested that the mutation in some way lowered the phosphatase activity of the protein and consequently, impaired proper transcription termination (Loya et al., 2012).

The SSU72 TOV mutation is of interest due to the fact that it implies an unknown function of the phosphatase or reinforces a previously identified function. The TOV mutant is also intriguing since the leucine to phenylalanine substitution in Ssu72 potentially highlights the importance of a region in the protein that is downstream of the identified catalytic site (highlighted in Figure 3). Research presented herein seeks to explain the link between Ssu72 L84F and the observed TOV phenotype. Specifically, this work aims to show that the TOV phenotype is potentially the result of an increase in phosphatase activity of the L84F mutant Ssu72 when compared to wild type Ssu72 in assays using a small molecule substrate. The result of this increase in activity results in higher levels of de-phosphorylated RNA Pol II CTD with consequential lower association of termination factors (NNS complex) with the CTD of Pol II. The findings shown provide insight into the function of an essential phosphatase that plays a large number of roles in transcription.

MATERIALS AND METHODS

I. Transformation of Wild Type plasmid DNA into *E. coli*

Wild type SSU72 DNA (pGST-SSU72 wild type supplied by the Faye lab) was amplified by adding 1 microliter of dsDNA to a vial of One Shot® chemically competent E.coli cells. After gentle mixing, the mixture was incubated for five minutes at 2-8°C. The sample was then mixed with in an Eppendorf thermomixer at 42°C for 30 seconds. 250 µL of S.O.C. Medium (Super Optimal Broth with Catabolite repression, Invitrogen part number 46-0700, lot 1141705) was then added to the tube and put into an incubator at 37°C, shaking at 150 rpm (Innova 44 incubator). After approximately 1 hour of shaking, 100 microliters of the transformation mix was transferred to a pre-warmed agar plate (LB + ampicillin). Culture was spread across the plate and the plate was incubated overnight at ambient conditions. A single colony was inoculated into LB with ampicillin (0.7X). The culture was incubated at 37°C with shaking (150 rpm). The next day, the culture tube was removed from the incubator and the broth was centrifuged at 6800 rcf for 2 minutes (Eppendorf 5430R centrifuge).

Plasmid DNA was isolated from the cells using a GeneJET Plasmid Miniprep Kit (part number K0503). Specifically, the cell pellet was re-suspended in 250 microliters of “Resuspension Solution” (with RNase A added). The solution was transferred to a centrifuge tube and 250 microliters of “Lysis Solution” was added and the sample was mixed. Next 350 microliters of “Neutralization Solution” was added to the sample and the mixture was clarified by centrifugation at 12000 rcf (at a higher rate for 5 minute. The supernatant was transferred to a GeneJET spin column, and 500 microliters of “Wash Solution” (diluted with ethanol prior to use) addition to the GeneJET column. The column was subjected to centrifugation at ~ 12,000 rcf (relative centrifugal force) or 30-60

seconds and the flow through was discarded. The wash step was repeated a second time and the tube was centrifuged an additional 1 minute to remove residual liquid. The spin column was transferred to a new centrifuge tube and 50 microliters of “Elution Buffer” was added to the column and incubated for 2 minute at room temperature followed by centrifugation for 2 minutes. “Elution Buffer” (50 microliters) was added to the column to elute more of the DNA. The plasmid DNA was retained for further use (referred to as p-SSU72 WT).

II. Site Directed Mutagenesis

An Agilent Quickchange Lightning Kit was used to conduct site directed mutagenesis of SSU72. Reactions were set up by combining 5 microliters of 10X “Reaction Buffer”, 10 – 100 ng of dsDNA template (SSU72 WT DNA), forward primer (Table 1 and Table 2), reverse primer (Table 1 and Table 2), 1 microliter of dNTP mix, 1.5 microliters of QuikSolution reagent and high purity water to a final volume of 50 microliters. 1 microliter of “QuickChange Lightning Enzyme” was then added to the mix.

Table 1. Sequence of primers used in site directed mutagenesis of GST tagged constructs

Primer	Sequence
SSU72_L84A_FORWARD	5’-ACGATCCAGCATTTGCAATGCACCGTTC-3’
SSU72_L84A_REV	5’-CCGTTACAAGTCGAACGGTGCATTGCAA-3’
SSU72_L84F_FORWARD	5’-CGATCCAGCATTTGCAAGAAACCGTTCG-3’
SSU72_L84F_REV	5’-GTTACAAGTCGAACGGTTTCTTGCAAATG-3’
ssu72 c15s F	5’-ATTGTTGTTTGATGCAGAAACTGTGCAAAACTTCAAGTTTGAATTGCG-3’
ssu72 c15s R	5’-CGCATTTCAAACTTGAAGTTTTGCACAGTTTCTGCATCAAACAACAAT-3’

Table 2. Details of primers used in site directed mutagenesis of GST tagged constructs

Primer	Primer Molecular Weight	Tm (°C)	Reconstituted Solution Concentration (mg/mL)	Mass added to Reaction (ng)
SSU72_L84A_FORWARD	12,256.0	68.0	1.16 mg/mL	125
SSU72_L84A_REV	12,336.0	68.0	1.24 mg/mL	124
SSU72_L84F_FORWARD	11,317.4	64.4	1.33 mg/mL	125
SSU72_L84F_REV	11,419.4	64.4	1.16 mg/mL	125
ssu72 c15s F	14,850.7	65.1	1.49 mg/mL	123
ssu72 c15s R	14,677.6	65.1	1.47 mg/mL	125

Reactions were cycled in a 96 well plate (Thermo Scientific AB3396) for the L84A and L84F constructs. The C15S construct was cycled in PCR tubes using the same thermocycler. Thermocycling was completed using an Eppendorf Mastercycler® Pro thermocycler with the parameters detailed below in Table 3.

Table 3. Thermocycler program for site directed mutagenesis

Segment	Cycles	Temperature (°C)	Time
1	1	95	2 minutes
2	18	95	20 seconds
		60	10 seconds
		68	2 min, 48 sec
3	1	68	5 minutes

Identities of the final plasmids are described below in Table 4.

Table 4. Plasmid construct descriptions

Construct	**Plasmid Identity**
SSU72 C15S	p-SSU72 C15S
SSU72 L84F	p-SSU72 L84F
SSU72 L84A	p-SSU72 L84A

III. Transformation of Rosetta Cells with SSU72 plasmids and glycerol stock preparation

Rosetta 2(DES) pLys(s) cells (20 microliters) were transferred to a 1.5 mL Eppendorf tube and the applicable plasmid DNA (per Table 5 below) was dispensed into the tube with the cells. The mixture was incubated on ice for 5 minutes and the tube was heat shocked in a 42°C for 30 seconds and then placed back into ice for 2 minutes. S.O.C. Medium (80 microliters) was added to the tube and 50 microliters of the cells were transferred to a pre-warmed agar plates with chloramphenicol and ampicillin antibiotic (spread by sterile glass beads). Plates were allowed to incubate at room temperature overnight. A culture tube with 2 mL of LB, 1X ampicillin and 1X chloramphenicol and the tube was inoculated with a single colony from the plate that had been incubated. The tube was incubated at 37°C with agitation at 150 rpm (Innova 44 incubator). Glycerol stocks of the cell lines were prepared and stored at -80°C. A table of the stocks prepared is shown below in Table 5.

Table 5. Details of transformation for each construct

Stock ID	SSU72 Construct	Source Plasmid	Mass Used in Transformation (ng)
pAM305	GST-SSU72 L84F	p-SSU72 L84F	55.3
pAM306	GST-SSU72 L84A	p-SSU72 L84A	51.6
pAM311	GST-SSU72 WT	p-SSU72 WT	56.0
pAM312	GST-SSU72 C15S	p-SSU72 C15S	101.1

IV. DNA Sequencing

The sequence of the mutated DNA was confirmed through third-party Sanger Sequencing (Genewiz). The details of these submissions along with the sequencing primer used with the samples are described in Table 6 and Table 7 below.

Table 6. Details of Genewiz Sequencing

Stock ID	SSU72 Construct	Genewiz Tracking Number
pAM305	GST-SSU72 L84F	10-309331369
pAM306	GST-SSU72 L84A	10-309331369
pAM311	GST-SSU72 WT	10-308769049
pAM312	GST-SSU72 C15S	10-311165623

Table 7. Genewiz sequencing primer details

Primer Number	Sequence Name	Sequence	Tm °C
P15-56	GST_544 TO 563	CGT ATT GAA GCT ATC CCA CA	52

V. GST-SSU72 Expression in *Escherichia coli* and Affinity Purification

Starter cultures were generated by inoculating a small volume of LB (i.e. 1 – 5 mL), 1X ampicillin/1X chloramphenicol with glycerol stock from the applicable construct (i.e. pAM305, pAM306, pAM311, pAM312). Starter culture was incubated several hours to overnight at 37°C and used to inoculate 1L of the same LB/antibiotic broth (in a 2L Erlenmeyer flask). Culture was grown to OD600 of at least 0.8 followed by addition of 1 mM IPTG at which point the incubation temperature was shifted to 18°C. After overnight protein expression, cells in the 1L broth were pelleted (Thermo Scientific Sorvall RC6+ centrifuge with rotor F9-4x1000y) and the pellet was re-suspended in lysis buffer (50 mM HEPES, 150 mM NaCl, 10% Glycerol, 10 mM EDTA, 1 mM β-mercaptoethanol) supplemented with 1 mM PMSF and 0.1% Triton X-100. The cells were then transferred to a chilled pressure cell that was set in a French Press manifold and cells were lysed three times at a pressure of ~ 1000 psi (high setting). Lysates were clarified by centrifugation using a Thermo Scientific Sorvall RC6+ centrifuge (with rotor F13-14x50cy) at 13,000 rpm for 60 minutes and the supernatant liquid was loaded onto a Biorad column packed with 1 mL of Glutathione Sepharose 4B resin (GE Healthcare). The loading occurred at a rate of 1 mL/min using the Biorad chromatography system. After loading, the column was washed with a wash buffer solution (50 mM HEPES, 300 mM NaCl, 10% Glycerol, 10 mM EDTA and 1 mM β-mercaptoethanol). Specifically a

gradient of lysis buffer to 100% wash buffer was used at a rate of 1 mL/min for 30 minutes (lysis buffer to wash buffer). After 30 minutes, the column was washed with 100% wash buffer for another 30 minutes. After washing, the GST tagged protein was eluted from the column using an elution buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM EDTA, 0.1-1 mM β-mercaptoethanol, 100 mM R-glutathione). Elution was conducted using an increasing gradient of elution buffer at a rate of 1 mL/min (lysis buffer to elution buffer). Elution fractions (1 mL each) were collected each minute as the UV absorbance was recorded. Fractions were combined based on the recorded absorbance and the combined fractions were transferred to a 7000 MWCO dialysis tube. Dialysis occurred overnight at 2-8°C in 5L of dialysis buffer (50 mM HEPES, 150 mM NaCl, 1 mM DTT). The next day, the dialysate was transferred to a 10 kD spin filter and the solution was concentrated to around 1 mL. Glycerol was added to the stock (around 10% of the total volume) and the stock was stored at -80°C for future use or analysis.

VI. Protein Analysis by SDS-PAGE and Coomassie Stain

SDS-PAGE gels were prepared per the recipe shown below in Table 8. Gels were loaded with Biorad Precision dual color standard and samples to be analyzed were diluted as desired with SDS running buffer and heat denatured at 100°C for approximately 5 minutes. A BSA protein standard was also used and was loaded into appropriate wells in the gel for estimation of protein concentrations.

Table 8. SDS-PAGE gel formulation

Component	**Stacking**	**Component**	**Separating**
30% Polyacrylamide (mL)	1.36	30% Polyacrylamide (mL)	10
1M Tris (pH 6.8) (mL)	1	1M Tris (pH 8.8) (mL)	7.5
10% Ammonium persulfate (mL)	0.08	10% Ammonium persulfate (mL)	0.2
20% SDS (mL)	0.04	20% SDS (mL)	0.1
TEMED (mL)	0.008	TEMED (mL)	0.008
Water (mL)	5.44	Water (mL)	4.6
Total (mL)	7.928	Total (mL)	22.408
Gel % SDS	5	Gel % SDS	13

Electrophoresis was conducted using a SDS gel box and 1X TGS buffer. Gels electrophoresis for 45 minutes at 200 volts. Proteins separated in the gels were stained with Coomassie stain and de-stained overnight with a solution containing 7% methanol and 6% glacial acetic acid. Stained proteins were visualized using an Epson V700 scanner and gel concentrations were estimated using the ImageJ open source program.

VII. GST Tag Removal

Purified GST-Ssu72 WT construct were thawed along with GST-3C prepared by Asha Boyd on 8/28/15. Purified WT construct (375 microliters) was added to a 1.5 mL Eppendorf tube. Approximately 1000 microliters of chilled cleavage buffer solution (10

mM Tris-HCl pH 8, 150 mM NaCl, 0.1% IGEPAL CA-630, and 0.5 mM EDTA (pH 8)) was added to the Ssu72 WT construct along with 10 microliters of the GST-3C protease. The sample was agitated overnight via rotisserie.

A Biorad column was packed with 1 mL of Glutathione Sepharose 4B beads and cleavage buffer was used in this process of preparing the column for use. The column was washed with water purified by a Sartorius system followed by wash with cleavage buffer. The column was loaded with the cleaved Ssu72 recombinant protein at a flow rate of 1 mL/min. Cleavage protein was collected during this rinse. The GST was then eluted from the column through an elution buffer solution of 150 mM NaCl, 40 mM EDTA, 50 mM HEPES, and 100 mM reduced L-glutathione adjusted to pH 7.6. The protein was collected and concentrated in a 10kD spin filter prior to further analysis.

VIII. *In-Vitro* Ssu72 Phosphatase Activity Assay

The phosphatase assay was completed in a Costar black 96 well plate. Each well was filled with Molecular Biology grade water to achieve a final volume of 100 - 350 µL depending on the experiment, 250 mM succinic acid, pH 6.0 (to a final concentration of 50 mM), 5M NaCl (to a final concentration of 150 mM NaCl), dialysis buffer solution (per the composition described in the expression section) sufficient to achieve equivalent volume of Ssu72 protein solution added to each well in the case that multiple constructs is tested, Ssu72 protein and a final concentration of 10 µM 6,8-difluoro-4-methylumbelliferyl phosphate (DiFMUP) in DMSO. After addition of DiFMUP, the plate was analyzed in a Molecular Devices Spectramax® M5 plate reader set at 30°C with excitation at 355 nm and emission at 455 nm.

IX. Biotinylated CTD peptide pulldown experiments

The pulldown protocol was adapted from the procedure from the manufacturer of the Dynabeads® (Invitrogen). 2.5 µg of biotinylated CTD peptide (four heptad repeats) was incubated with 0.5 mg of Invitrogen Dynabeads® M280 Streptavidin in binding buffer (50 mM Tris-HCl pH 6.5, 300mM NaCl, 1 mM dithiothreitol, 0.5% IGEPAL CA-630, 1mM PMSF) at 4 °C for 2 hours on a rotisserie. Recombinant GST-Ssu72 (approximately 10 µg) and 100 µg of BSA to the micro centrifuge tube were then added to the respective tubes and incubated for another 3 hours at 4°C. Samples were then inserted into a magnetic separation rack. Beads were given ample time to separate from the supernatant and the liquid was removed via pipette. The beads were then washed four times with 1 ml binding buffer containing 10 µg/mL BSA and in each case, previously mentioned magnetic separation and liquid removal methods were used. Post wash, 100 µl of 4x SDS-loading buffer was added to each of the tubes and each were incubated in a heat block at 100°C for 5 minutes to separate biotinylated protein from Streptavidin beads. Beads were separated from the liquid using the magnetic rack and the supernatant liquid was retained for analysis.

X. Expression and Purification of Ssu72-3xFLAG in *Saccharomyces cerevisiae*

Log-phase yeast cells expressing Ssu73-3XFLAG were collected by centrifugation using rotor the RC6+ floor centrifuge at 4000 x g for 10 minutes. The cell pellet was washed with high purity water and re-suspended in 25mL of TAP lysis buffer (40mM HEPES-KOH, pH 7.5, 10% glycerol, 350mM NaCl, 0.1% Tween-20, 1X fresh yeast protease inhibitors solution per 2.5 grams of pellet). Cells were lysed by grinding in a Waring blender with dry ice after freezing in liquid nitrogen. Lysates were thawed and the resulting extract was treated with 100 units of DNAse I and 10 microliters of 30 mg/mL

heparin for 10 minutes at room temperature. The extract was then centrifuged at 14,000 x g for 1 hour in 50mL conical tubes in the RC6+ centrifuge to pellet the remaining debris (sample remaining in supernatant). Anti-FLAG agarose resin (500 µL) was washed with 1.5 mL TAP lysis buffer in a micro centrifuge tube and the resin was re-suspended in 500uL TAP lysis buffer. The 50mL conical tubes were removed from the centrifuge once centrifugation was complete and clarified lysate was decanted to a clean vessel. Washed anti-FLAG agarose (Sigma) was added to the lysate and incubated on a stir plate overnight at 4°C. After an overnight incubation, the cell suspension was transferred to a 30mL Bio-Rad Econoprep column and drained by gravity flow. Beads retained in the column were then washed with 120 mL TAP lysis buffer solution by gravity flow. FLAG peptide (125 µL) was added to resin in the column followed by incubation for 5 minutes (using a cap on the outlet of the column to prevent draining). Liquid in the column was then eluted by gravity. 125 uL of TAP lysis buffer was used to wash the resin (combined wash and elution were labeled as elution 1). Elution and wash steps were then repeated five times and eluted aliquots were stored at -80°C until needed. Each elution after elution 1 was labeled sequentially E2 through E6.

XI. Protein Analysis by SDS-PAGE and Silver Stain

Samples from the yeast expression and purification study were analyzed by SDS-PAGE followed by silver staining. Purified samples were first diluted 1:1 with SDS-PAGE running buffer (20 microliters of elution sample with 20 microliters of 2X SDS Loading Buffer) followed by denaturation at 100 °C for 5 minutes. Denatured samples in a volume of 20 µL were loaded into a Bio-Rad Mini Protean® precast gel and electrophoresis was performed in a SDS polyacrylamide gel box filled with 1X TGS buffer at 200 volts for 35 minutes. Following electrophoresis, gels were removed from the assembly and covered

with ~100mL of fixing solution at room temperature overnight. Fixing solution was discarded and ~ 100mL of ethanol wash was added to the container at room temperature with rocking for around 10 minutes. Ethanol wash was discarded and ~100mL of Sartorius water was added with rocking for about 10 minutes. Water was discarded and ~100mL of Sensitizer Solution was added, followed by rocking for ~10 minutes. Sensitizer solution was discarded and ~100mL of Sartorius water was at room temperature for 10 minutes. Water wash was discarded and ~100mL of silver nitrate solution was added with rocking for ~10 minutes. Silver nitrate solution was discarded with ~ 100 mL of Sartorius water added to the container. The gel was washed with rocking for ~5 minutes. Water was discarded and ~100mL of fresh developing solution was added with mild agitation for several minutes until the desired development of the gel was reached. Developing solution was discarded at this point and the gel staining was quenched using stop solution for 10 minutes. Stop solution was then removed and the gel was stored in water until imaged for documentation. Formulations for the solutions used in this work are summarized below in Table 9.

Table 9. Silver Stain Stock Formulations

Stock	Composition
Fixing Solution	30% Ethanol 10% Acetic Acid 60% Water
Ethanol Wash	30% Ethanol 70% Water
Sensitizer Solution	0.02% Sodium Thiosulfate (2% Solution) 99.98% Water
Silver Nitrate Solution	0.1% Silver Nitrate 0.02% Formaldehyde 99.9% Water
Developing Solution	2.5% Sodium Carbonate 0.05% Formaldehyde 0.005% Sodium Thiosulfate 97.445% Water
Stop Solution	0.5% Glycine 99.5% Water

XII. Analysis of 3xFLAG WT Ssu72 by LC-MS/MS

These experiments were performed as previously described (Smith-Kinnaman et al., 2014). Two of the elutions (2 and 3) from the *S. cerevisiae* expression and purification in section X (100 µL of each) were combined in a micro centrifuge tube and 200 µl of cold solution of 100 mM Tris-HCl (pH 8.5) was added. Trichloroacetic Acid (100 µL) was then

added to the tube and mixed overnight at 4°C to precipitate the protein. The pellet was collected by centrifugation at 14,000 rpm for 30 min at 4°C (Thermo Scientific Sorvall RC6+ centrifuge with rotor F13-14x50cy). The supernatant was discarded and 500 µl of cold acetone was added to the pellet and mixed by vortexing. The pellet was collected by centrifugation for 10 min at 4°C at 14,000 rpm; the acetone wash was repeated once more in order to remove the TCA. After the second wash was decanted, the tube was left open to allow acetone to evaporate from the pellet. After precipitation was complete, the sample was ready for digestion. First 30 µl of solution of 8M Urea in 100 mM Tris-HCl (pH 8.5) was added to the TCA-precipitant. A 1.5 µL aliquot of 1 M Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) was then added to the tube to achieve a final concentration of 50 mM TCEP. This step was conducted in order to reduce the disulfide bonds of the protein. After TCEP addition, the contents of the mixture was incubated at room temperature for 30 minutes. A 0.6 µL aliquot of 0.5M chloroacetamide (CAM) was added to the reduced mixture to a final concentration of 10mM CAM, and incubated at room temperature for 30 min in the dark. The CAM addition was used to alkylate reduced bonds in order to prevent oxidation. Endoproteinase Lys-C (0.3 µL of 0.2µg/µl solution in water) was added to the alkylated mixture and incubated at 37°C with shaking overnight. Post proteolytic cleavage, 90 µL of 100mM Tris-HCl, pH 8.5 was added to the mixture to dilute Urea to 2M followed by addition of 0.6 µL 1M CaCl 2 (final concentration 2mM). Trypsin (0.5 µg) was added with further overnight incubation at 37°C and the cleavage reaction was quenched the next day with 7 µL formic acid.

Column loading was conducted using a procedure described previously (Florens & Washburn, 2006). A fused silica glass column was packed with two different resins via a high pressure vessel and resin suspended in methanol. Packing consisted of 8.5 cm of

reverse phase C18 (Phenomonex Aqua®), 2.5 cm of strong cation exchange resin (Phenomonex Luna®) and 2.25 cm of C18. Post packing, the column was rinsed with buffer A solution (5% acetonitrile and 0.1% formic acid) and loaded with the digested/quenched Ssu72 sample. The column was then rinsed again with buffer A solution after loading in preparation for analysis. The Ssu72 protein was eluted from the column using a Proxeon nano-liquid chromatograph (LC). The LC was in-line with a ThermoFisher LTQ Velos linear ion trap mass spectrometer (MS). Elution from the column was analyzed using a 10-step multi-dimensional protein identification technology (MudPIT) method. Separation of components in the loaded column was accomplished by a gradient elution from buffer A (5% acetonitrile, 0.1% formic acid) to buffer B (80% acetonitrile, 0.1% formic acid). Mass spectra were analyzed with Scaffold (version 4) where Steps 2 through 4 of the data set were queried further while using a 5.0% peptide false discovery rate (FDR), 95% protein threshold and 2 minimum peptides to identify.

RESULTS

I. Sequencing

The sequence of the four GST-SSU72 constructs was confirmed by Sanger Sequencing (Genewiz) in order to confirm that the plasmid mutations were correct for each of the constructs. The results reported by Genewiz aligned with SSU72 (S288C) are shown below in Figure 5. The results confirmed that the desired mutations were achieved by site directed mutagenesis.

```
CLUSTAL O (1.2.1) multiple sequence alignment


SSU72           ------------------------ATGCCTAGTCATCGCAATTCAAACTTGAAGTTTTGCA
pAM311-WT       CCCTGGGATCCCCGGAATTCCGGATGCCTAGTCATCGCAATTCAAACTTGAAGTTTTGCA
pAM312-C15S     CCCTGGGATCCCCGGAATTCCGGATGCCTAGTCATCGCAATTCAAACTTGAAGTTTTGCA
pAM306-L84A     CCCTGGGATCCCCGGAATTCCGGATGCCTAGTCATCGCAATTCAAACTTGAAGTTTTGCA
pAM305-L84F     CCCTGGGATCCCCGGAATTCCGGATGCCTAGTCATCGCAATTCAAACTTGAAGTTTTGCA
                                         ************************************


SSU72           CAGTTTGTGCATCAAACAACAATCGTTCAATGGAATCGCATAAAGTCCTGCAAGAAGCAG
pAM311-WT       CAGTTTGTGCATCAAACAACAATCGTTCAATGGAATCGCATAAAGTCCTGCAAGAAGCAG
pAM312-C15S     CAGTTTCTGCATCAAACAACAATCGTTCAATGGAATCGCATAAAGTCCTGCAAGAAGCAG
pAM306-L84A     CAGTTTGTGCATCAAACAACAATCGTTCAATGGAATCGCATAAAGTCCTGCAAGAAGCAG
pAM305-L84F     CAGTTTGTGCATCAAACAACAATCGTTCAATGGAATCGCATAAAGTCCTGCAAGAAGCAG
                ****** **************************************************


SSU72           GCTATAATGTTAGCTCTTACGGAACAGGTTCAGCTGTGAGACTGCCTGGTCTATCGATAG
pAM311-WT       GCTATAATGTTAGCTCTTACGGAACAGGTTCAGCTGTGAGACTGCCTGGTCTATCGATAG
pAM312-C15S     GCTATAATGTTAGCTCTTACGGAACAGGTTCAGCTGTGAGACTGCCTGGTCTATCGATAG
pAM306-L84A     GCTATAATGTTAGCTCTTACGGAACAGGTTCAGCTGTGAGACTGCCTGGTCTATCGATAG
pAM305-L84F     GCTATAATGTTAGCTCTTACGGAACAGGTTCAGCTGTGAGACTGCCTGGTCTATCGATAG
                ************************************************************
```

```
SSU72           ATAAGCCTAATGTGTACTCATTTGGTACACCCTATAATGATATATATAATGATCTTTTAT
pAM311-WT       ATAAGCCTAATGTGTACTCATTTGGTACACCCTATAATGATATATATAATGATCTTTTAT
pAM312-C15S     ATAAGCCTAATGTGTACTCATTTGGTACACCCTATAATGATATATATAATGATCTTTTAT
pAM306-L84A     ATAAGCCTAATGTGTACTCATTTGGTACACCCTATAATGATATATATAATGATCTTTTAT
pAM305-L84F     ATAAGCCTAATGTGTACTCATTTGGTACACCCTATAATGATATATATAATGATCTTTTAT
                ************************************************************

SSU72           CACAATCAGCAGACCGTTACAAGTCGAACGGTTTATTGCAAATGCTGGATCGTAATAGAA
pAM311-WT       CACAATCAGCAGACCGTTACAAGTCGAACGGTTTATTGCAAATGCTGGATCGTAATAGAA
pAM312-C15S     CACAATCAGCAGACCGTTACAAGTCGAACGGTTTATTGCAAATGCTGGATCGTAATAGAA
pAM306-L84A     CACAATCAGCAGACCGTTACAAGTCGAACGGTGCATTGCAAATGCTGGATCGTAATAGAA
pAM305-L84F     CACAATCAGCAGACCGTTACAAGTCGAACGGTTTCTTGCAAATGCTGGATCGTAATAGAA
                ********************************   ************************

SSU72           GACTCAAAAAAGCACCTGAAAAATGGCAAGAAAGTACAAAAGTCTTCGACTTCGTTTTCA
pAM311-WT       GACTCAAAAAAGCACCTGAAAAATGGCAAGAAAGTACAAAAGTCTTCGACTTCGTTTTCA
pAM312-C15S     GACTCAAAAAAGCACCTGAAAAATGGCAAGAAAGTACAAAAGTCTTCGACTTCGTTTTCA
pAM306-L84A     GACTCAAAAAAGCACCTGAAAAATGGCAAGAAAGTACAAAAGTCTTCGACTTCGTTTTCA
pAM305-L84F     GACTCAAAAAAGCACCTGAAAAATGGCAAGAAAGTACAAAAGTCTTCGACTTCGTTTTCA
                ************************************************************

SSU72           CTTGTGAAGAGAGATGTTTTGATGCCGTTTGTGAAGATTTGATGAATAGAGGTGGGAAAT
pAM311-WT       CTTGTGAAGAGAGATGTTTTGATGCCGTTTGTGAAGATTTGATGAATAGAGGTGGGAAAT
pAM312-C15S     CTTGTGAAGAGAGATGTTTTGATGCCGTTTGTGAAGATTTGATGAATAGAGGTGGGAAAT
pAM306-L84A     CTTGTGAAGAGAGATGTTTTGATGCCGTTTGTGAAGATTTGATGAATAGAGGTGGGAAAT
pAM305-L84F     CTTGTGAAGAGAGATGTTTTGATGCCGTTTGTGAAGATTTGATGAATAGAGGTGGGAAAT
                ************************************************************

SSU72           TAAACAAAATAGTGCATGTAATTAATGTTGACATTAAAGATGATGATGAAAATGCTAAAA
pAM311-WT       TAAACAAAATAGTGCATGTAATTAATGTTGACATTAAAGATGATGATGAAAATGCTAAAA
pAM312-C15S     TAAACAAAATAGTGCATGTAATTAATGTTGACATTAAAGATGATGATGAAAATGCTAAAA
pAM306-L84A     TAAACAAAATAGTGCATGTAATTAATGTTGACATTAAAGATGATGATGAAAATGCTAAAA
pAM305-L84F     TAAACAAAATAGTGCATGTAATTAATGTTGACATTAAAGATGATGATGAAAATGCTAAAA
                ************************************************************
```

```
SSU72           TTGGTAGCAAAGCTATATTGGAATTAGCTGATATGCTCAATGATAAAATAGAACAATGTG
pAM311-WT       TTGGTAGCAAAGCTATATTGGAATTAGCTGATATGCTCAATGATAAAATAGAACAATGTG
pAM312-C15S     TTGGTAGCAAAGCTATATTGGAATTAGCTGATATGCTCAATGATAAAATAGAACAATGTG
pAM306-L84A     TTGGTAGCAAAGCTATATTGGAATTAGCTGATATGCTCAATGATAAAATAGAACAATGTG
pAM305-L84F     TTGGTAGCAAAGCTATATTGGAATTAGCTGATATGCTCAATGATAAAATAGAACAATGTG
                ************************************************************

SSU72           AAAAAGATGACATTCCCTTTGAAGATTGTATAATGGACATTTTAACTGAGTGGCAAAGCT
pAM311-WT       AAAAAGATGACATTCCCTTTGAAGATTGTATAATGGACATTTTAACTGAGTGGCAAAGCT
pAM312-C15S     AAAAAGATGACATTCCCTTTGAAGATTGTATAATGGACATTTTAACTGAGTGGCAAAGCT
pAM306-L84A     AAAAAGATGACATTCCCTTTGAAGATTGTATAATGGACATTTTAACTGAGTGGCAAAGCT
pAM305-L84F     AAAAAGATGACATTCCCTTTGAAGATTGTATAATGGACATTTTAACTGAGTGGCAAAGCT
                ************************************************************

SSU72           CACATTCTCAACTACCGTCATTATACGCTCCTTCATATTACTAA----------------
pAM311-WT       CACATTCTCAACTACCGTCATTATACGCTCCTTCATATTACTAACCGCTCGAGCGGCCGC
pAM312-C15S     CACATTCTCAACTACCGTCATTATACGCTCCTTCATATTACTAACCGCTCGAGCGGCCGC
pAM306-L84A     CACATTCTCAACTACCGTCATTATACGCTCCTTCATATTACTAACCGCTCGAGCGGCCGC
pAM305-L84F     CACATTCTCAACTACCGTCATTATACGCTCCTTCATATTACTAACCGCTCGAGCGGNCGC
                ******************************************
```

Figure 5. Nucleotide Sequence Alignment for Ssu72 Constructs

II. Expression and Purification of GST-Tagged Ssu72

Individual batches of LB broth were inoculated using glycerol stocks from pAM305 (L84F mutant), pAM306 (L84A mutant), pAM311 (WT) and pAM312 (C15S mutant) respectively. Starter cultures were grown followed by induction and expression. Cell lysis was conducted and the proteins were purified using a Biorad system. Absorbance over time at 280 nm is shown below for each of the four constructs (Figure 6, Figure 7, Figure 8, and Figure 10).

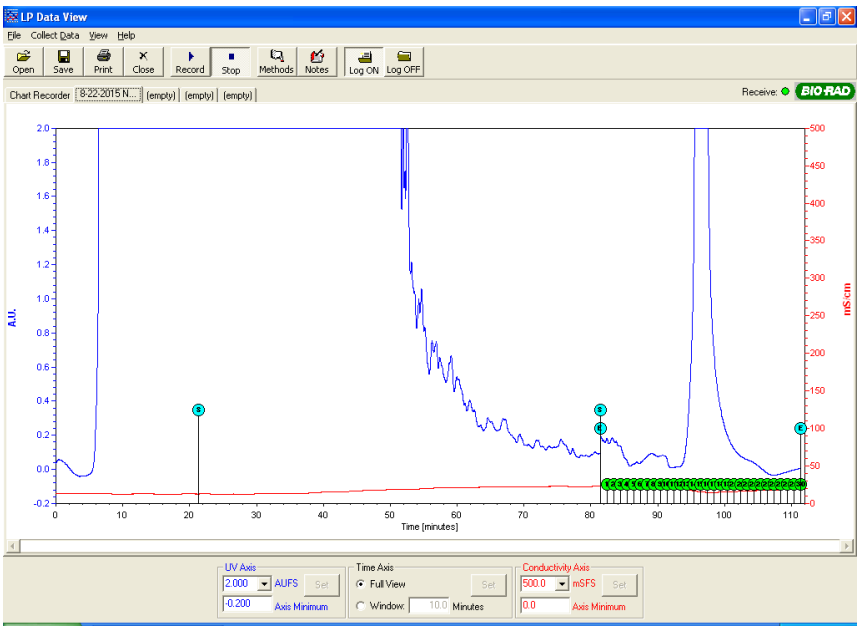


Figure 6. Chromatogram for Loading, Wash and Elution of GST-Tagged Ssu72 (Wild Type)

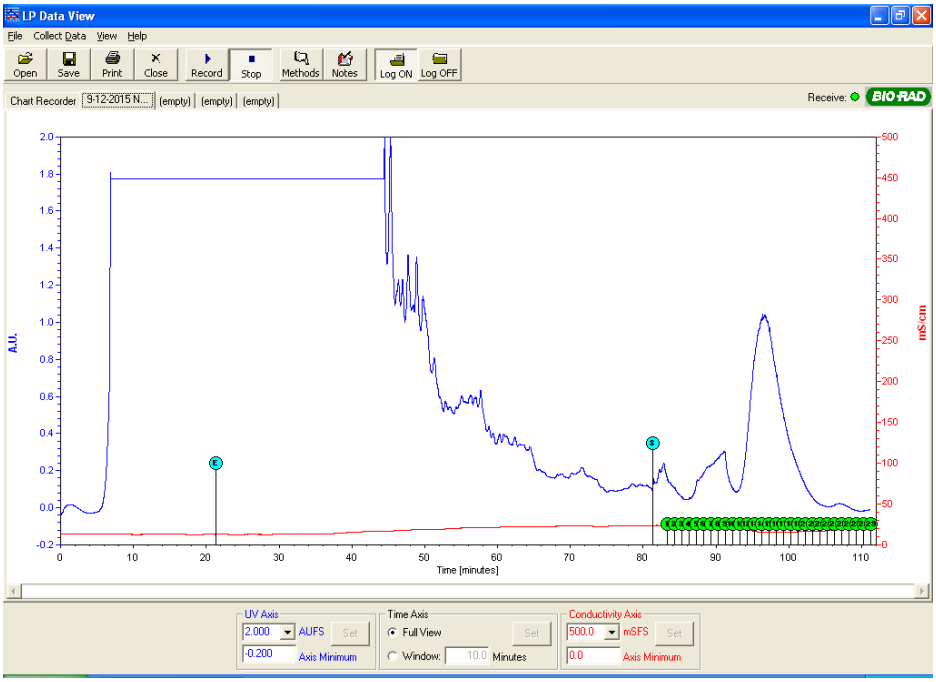


Figure 7. Chromatogram for Loading, Wash and Elution of GST-Tagged Ssu72 (C15S)

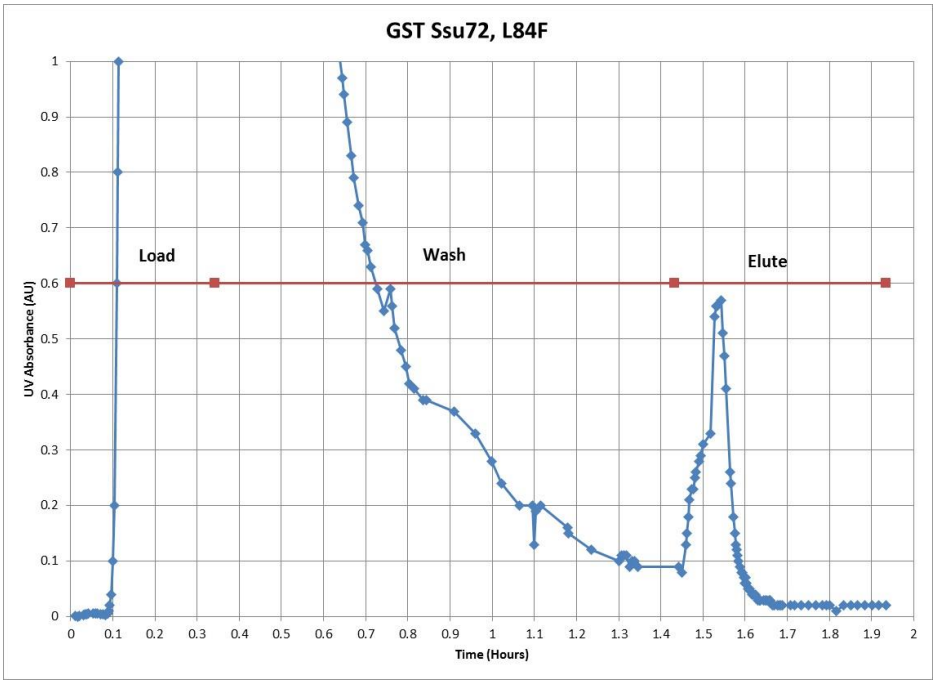


Figure 8. Chromatogram for Loading, Wash and Elution of GST-Tagged Ssu72 (L84F)

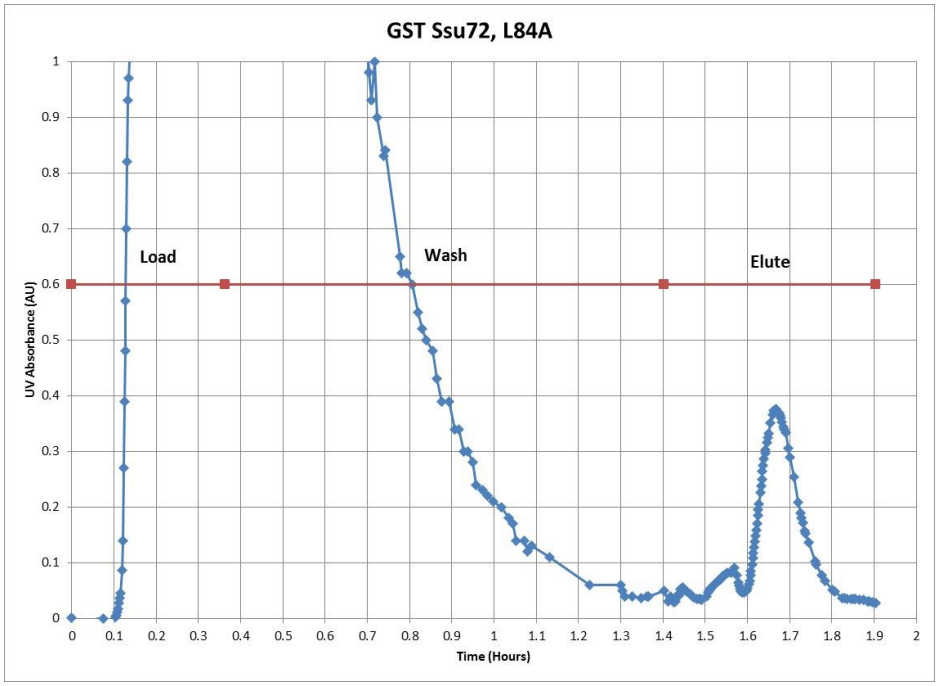
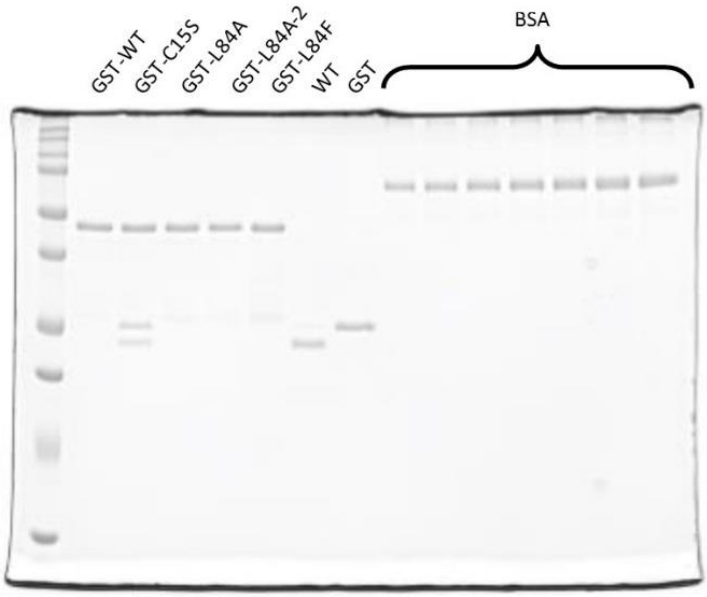


Figure 9. Chromatogram for Loading, Wash and Elution of GST-Tagged Ssu72 (L84A)

A total of six separate replicate purifications of protein constructs were prepared using this procedure for the later evaluation by DiFMUP-based phosphatase assays. These batches included WT Ssu72, C15S, two biological replicates of L84A and two biological replicates of L84F.

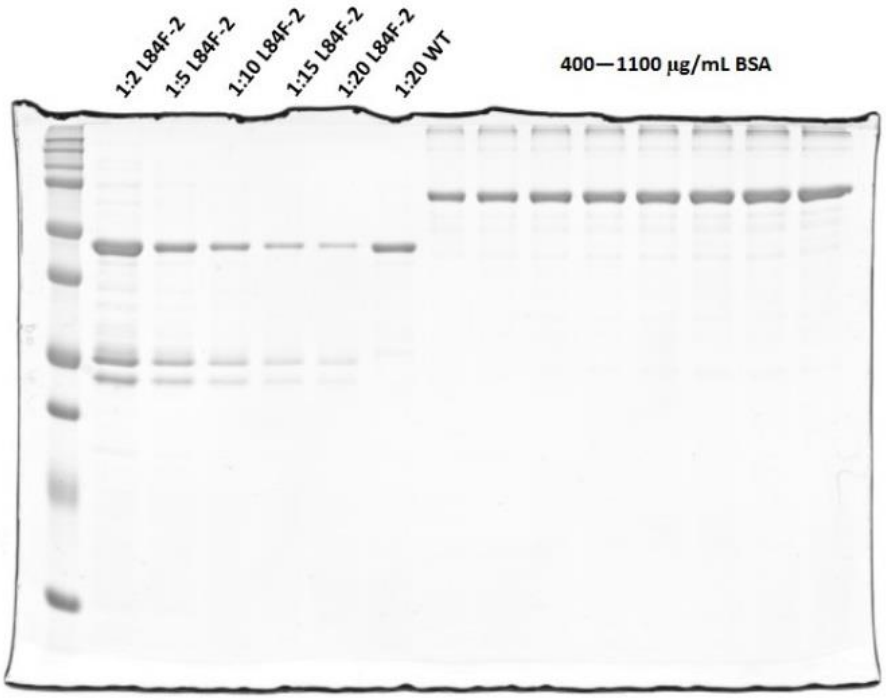
III. Analysis of GST Tagged Ssu72 Protein by SDS-PAGE and Coomassie Stain

The GST-Ssu72 constructs were analyzed by SDS-PAGE, followed by Coomassie staining. Bovine serum albumin was also analysis by gel electrophoresis at varying concentrations in order to estimate concentration of each purified Ssu72 sample. Proteins were visualized by silver staining, which were analyzed using ImageJ analysis software in order to approximate concentration of each purified sample. Gels along with the calculated concentrations of each recombinant Ssu72 protein were shown below in Figure 10 through Figure 12.



Protein	Dilution	Intensity	Protein Amount (µg)	Sample (mL)	Protein Concentration Adjusted for Dilution (mg/mL)
GST Ssu72 WT	20	12267.21	689.9	0.001	13.8
GST Ssu72 C15S	15	11803.08	657.1	0.001	9.9
GST Ssu72 L84A	15	11116.62	608.5	0.001	9.1
GST Ssu72 L84A2	10	9493.426	493.5	0.001	4.9
GST Ssu72 L84F	5	11328.43	623.5	0.001	3.1
Ssu72 WT	4	7455.719	349.3	0.001	1.4
GST	5	10453.79	561.5	0.001	2.8

Figure 10. SDS-PAGE, Coomassie stain of GST Ssu72 WT, C15S, L84A, L84A2, L84F, Cleaved WT, GST



Protein	Dilution	Intensity	Protein Amount (µg)	Sample (mL)	Protein Concentration Adjusted for Dilution (mg/mL)
L84F-2	2	35871.49	2855.7	0.001	5.7
L84F-2	5	18388.38	1262.3	0.001	6.3
L84F-2	10	10610.36	553.4	0.001	5.5
L84F-2	15	5654.61	101.7	0.001	1.5
L84F-2	20	3562.246	-89.0	0.001	-1.8
WT-2	20	9798.355	479.4	0.001	9.6

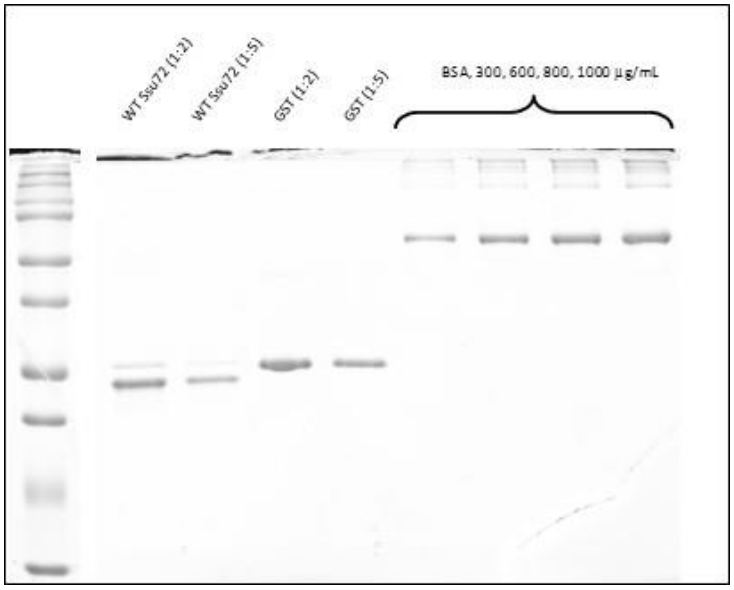
Figure 11. SDS-PAGE, Coomassie stain of GST Ssu72 L84F#2, WT (with smaller molecular weight bands)

Protein	Dilution	Intensity	Protein Amount (µg)	Sample (mL)	Protein Concentration Adjusted for Dilution (mg/mL)
L84F-2	2	11295.43	1127.6	0.001	2.3
L84F-2	5	6296.184	454.8	0.001	2.3
L84F-2	10	3883.527	130.1	0.001	1.3
L84F-2	15	2174.627	-99.9	0.001	-1.5
L84F-2	20	1396.92	-204.6	0.001	-4.1
WT-2	20	6506.648	483.1	0.001	9.7

Figure 12. SDS-PAGE, Coomassie stain calculations of GST Ssu72 L84F#2, WT (without smaller molecular weight bands)

IV. Removal of GST Tag from Ssu72 protein

In order to determine if the GST tag was impacting the activity of the phosphatase in the DiFMUP assay, the GST tag was removed from WT GST-Ssu72. In accomplish this removal, the purified GST-Ssu72 WT construct was first combined with GST-3C protease in cleavage buffer. After overnight cleavage, the cleaved protein was purified using a column packed with Glutathione Sepharose 4B beads. During the loading step of this purification, column was accidentally drained due to not shutting off the pump in time. Despite this issue, the column was rinsed with elution buffer and the eluted material was concentrated using a 10kD centrifuge filter. The purified protein was later analyzed by SDS-PAGE along with Coomassie staining to determine relative protein concentration. Results from this analysis are shown below in Figure 13 and demonstrate that the desired cleavage of the protein was achieved and that two proteins of different size were isolated. The slightly slower migrating bands in lanes 2 and 3 were confirmed to separate further by gel electrophoresis than the later migrating bands in lanes 4 and 5. Bands were near the 25 kD ladder band and were representative of the 23 kD Ssu72 and 26 kD GST.



Protein	Dilution	Intensity	Protein Amount (µg)	Protein Concentration Adjusted for Dilution (mg/mL)
Ssu72 WT	2	11092.21	957.0	1.9
Ssu72 WT	5	5139.477	312.1	1.6
GST	2	12606.13	1121.1	2.2
GST	5	7867.77	607.7	3.0

Figure 13. SDS-PAGE and Coomassie stain of Ssu72 and GST following GST tag cleavage

V. Assay of Phosphatase Activity *In-Vitro*

In the first experiment, relative fluorescence response was recorded for the four Ssu72 proteins (WT, C15S, L84A, L84A-2 and L84F). The Ssu72 proteins ranged in concentration from 0.01, 0.1, 1.0 and 10.0 µM and were analyzed in quadruplicate with DiFMUP at 10 µM (per conditions described in section VIII). In all experiments, WT Ssu72 was used as a positive control while C15S Ssu72 was used as a negative control. Fluorescence was measured by the plate reader and results are in Figure 14 through Figure 16 with assumed Ssu72 concentrations in Table 10. These data indicate that the L84A condition had a higher response than all other constructs when run at the 1 µM

concentration of Ssu72. While the response values for L84F had a higher average response than WT, L84F and WT responses for this experiment were within 1 standard deviation from each other. In the case of the 10 µM Ssu72 condition, the response was too rapid too show difference between most constructs but there was differentiation between L84F and wild type at the first two readings. C15S and GST were all lower than the other constructs.

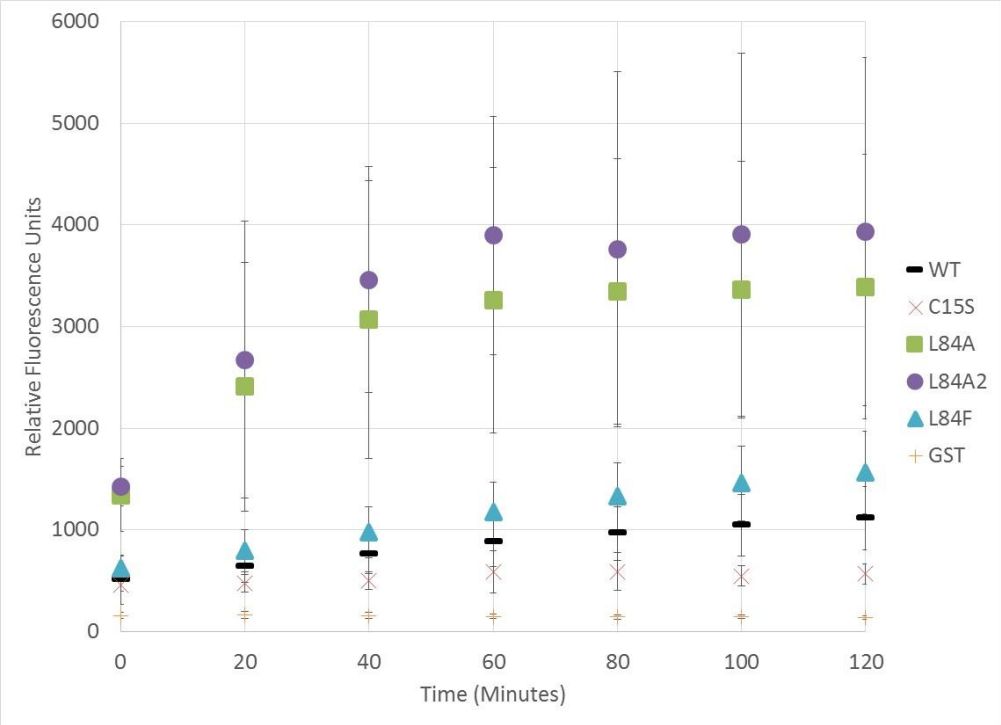


Figure 14. Kinetic study #1 with all constructs at 1 µM Ssu72 and DiFMUP at 10 µM (1 standard deviation shown)

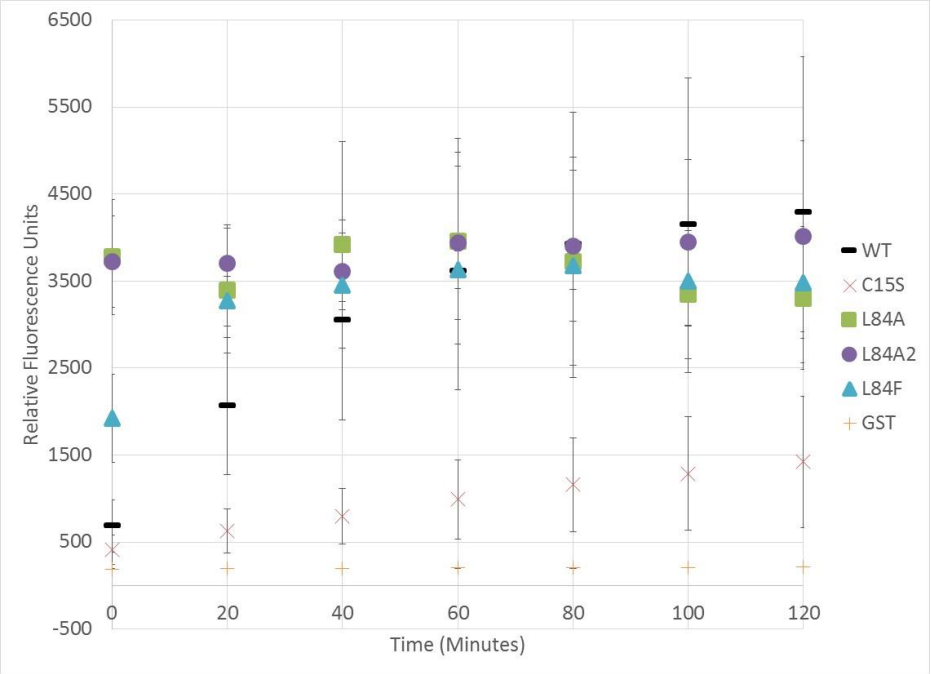


Figure 15. Kinetic study #1 with all constructs at 1 µM Ssu72 and DiFMUP at 10 µM (1 standard deviation shown)

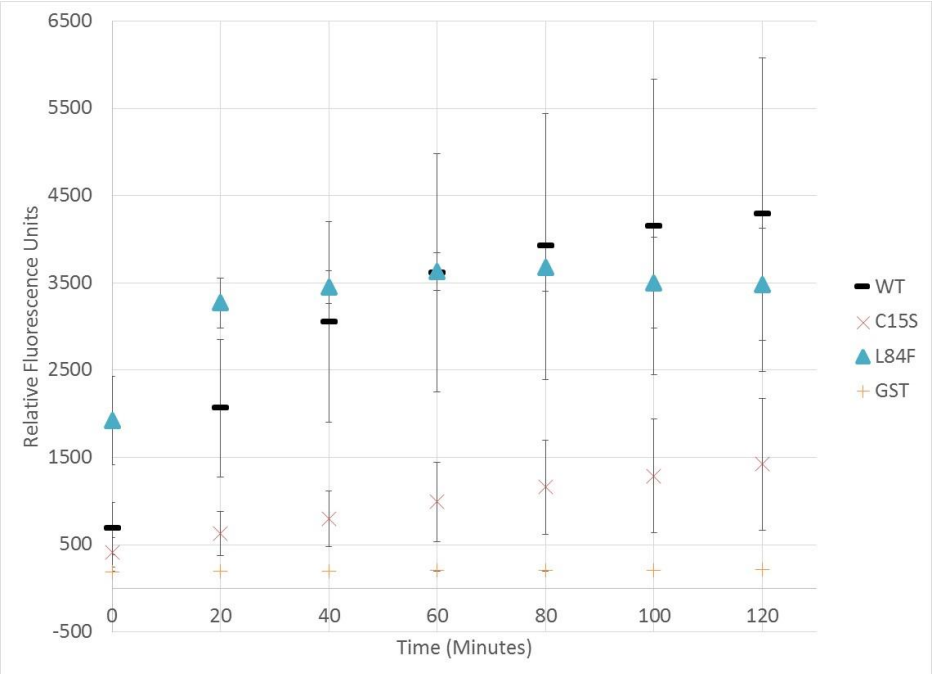


Figure 16. Kinetic study #1 with all constructs at 1 µM Ssu72 and DiFMUP at 10 µM, without L84A constructs (1 standard deviation shown)

Table 10. Assumed concentrations for kinetic study #1

Construct	Assumed Concentration (µg/µL)
GST-WT	13.8
GST-C15S	9.9
GST-L84A	9.1
GST-L84A2	4.9
GST-L84F	3.1

The second experiment kept DiFMUP concentration at 10 µM with Ssu72 concentrations of 0.25, 0.50, 1.0 and 2.0 µM. The time-course data for each of the constructs at the 2 µM condition is shown below in Figure 17. These results show a more definitive separation between L84F and WT.

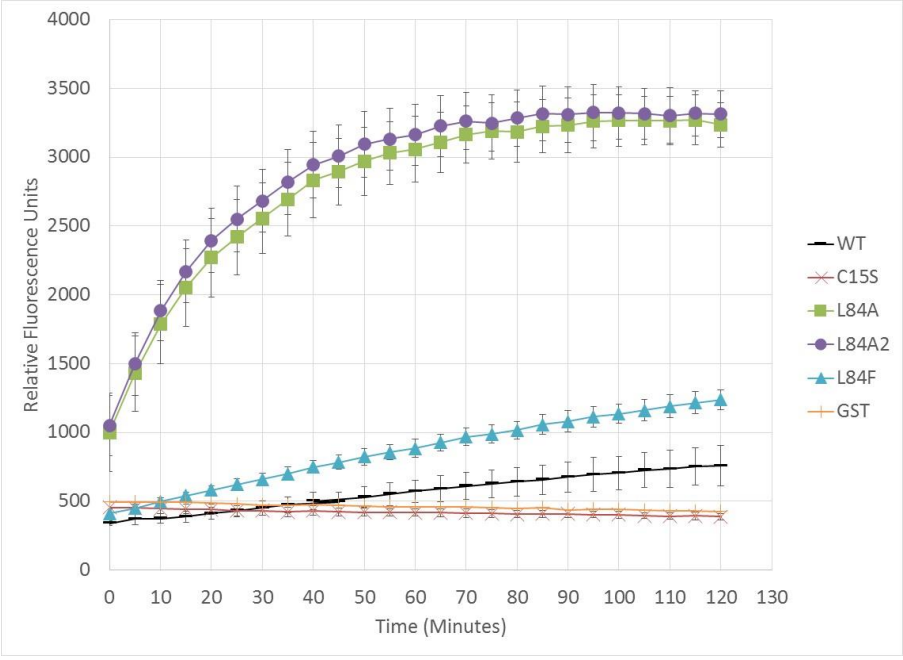


Figure 17. Kinetic study #2 with all constructs at 2 µM Ssu72 and DiFMUP at 10 µM (1 standard deviation shown)

The initial rate of change for each condition in this second experiment was calculated and the results are shown in Figure 18 and Figure 19. The assumed concentrations for the studies are shown in Table 11. These results reinforce the higher phosphatase activity for L84F over WT.

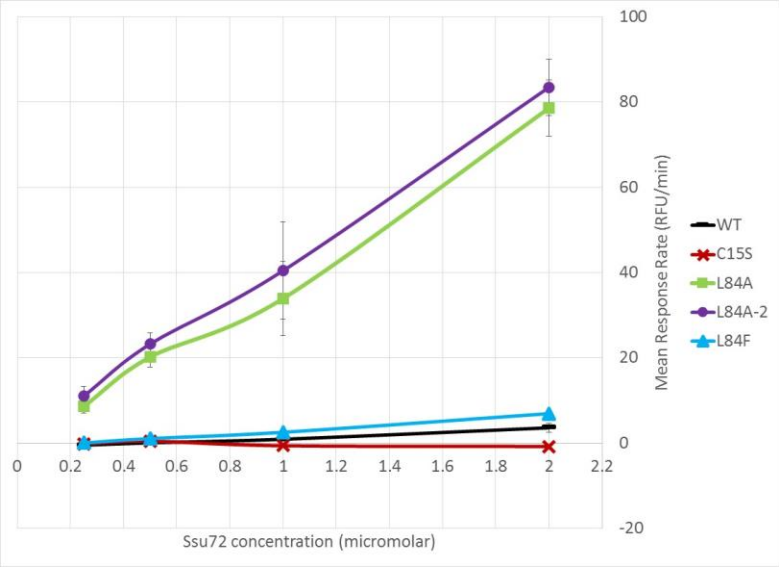


Figure 18. Kinetic study #2 0 to 2 µM (1 standard deviation shown)

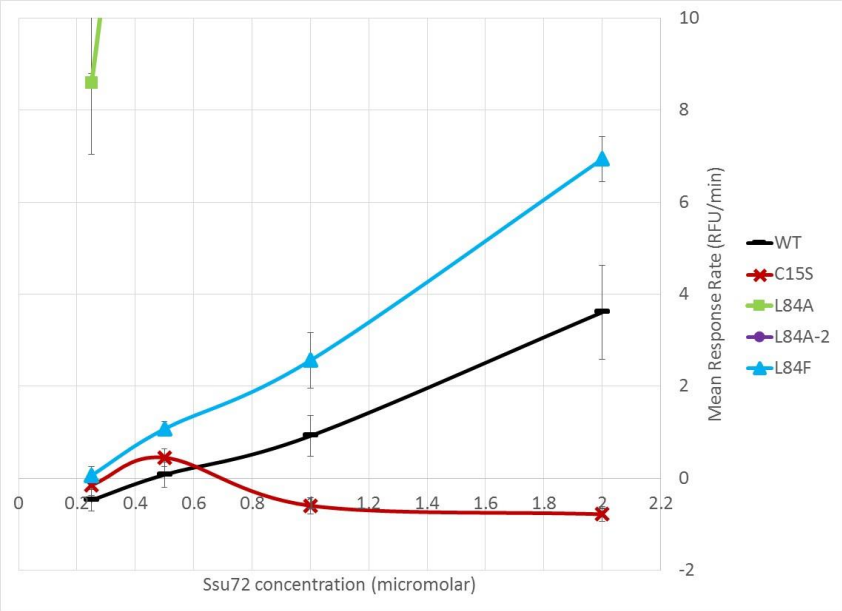


Figure 19. Kinetic study #2 0 to 2 µM (1 standard deviation shown), 0 to 2 µM, zoomed in (1 standard deviation shown)

Table 11. Assumed concentrations for kinetic study #2

Construct	Assumed Concentration (µg/µL)
GST-WT	13.8
GST-C15S	9.9
GST-L84A	9.1
GST-L84A2	4.9
GST-L84F	3.1

The data from the 2 µM kinetic study was also analyzed in order to address whether there was a difference between the cleaved Ssu72 and the GST tagged protein. The time course results from the replicates are shown in Figure 20.

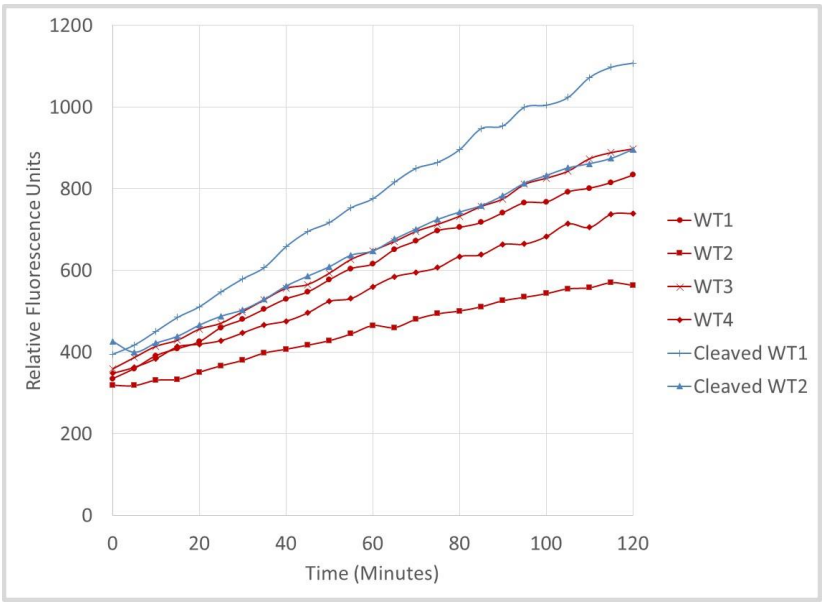


Figure 20. Relative Fluorescence vs. time for cleaved and GST tagged constructs at 2 µM

Data points were analyzed by Grubb's test to determine if each cleaved construct was an outlier when compared to the other GST-tagged constructs. More specifically, the final data point at 2 hours was used in the analysis due to the fact that the data points were furthest from the GST-tagged constructs in the data set and that this analysis would be worst case. As shown in Figure 21, G-crit is greater than G in both cases and consequently, the data points are not considered to be outliers when compared to the GST-tagged responses at 2 hours.

	Grubbs Test			
	Values		min	563.248
WT	833.552		mean	828.2526
WT	563.248		stdev	200.7201
WT	897.523		G	1.320269
WT	739.123			
C-WT	1107.817		alpha	0.05
			size	5
			sig value	0.01
			df	3
			t-crit	4.540703
			G-crit	1.671386
			sig value	No

	Grubbs Test			
	Values		min	563.248
WT	833.552		mean	785.9664
WT	563.248		stdev	140.2672
WT	897.523		G	1.587815
WT	739.123			
C-WT	896.386		alpha	0.05
			size	5
			sig value	0.01
			df	3
			t-crit	4.540703
			G-crit	1.671386
			sig value	No

Figure 21. Grubb's test analysis for cleaved Ssu72 (replicate #1 left, replicate #2 right)

The phosphatase activity of the Ssu72 constructs was also measured in the DiFMUP assay by varying substrate concentration while keeping phosphatase concentration constant at 2 µM. The DiFMUP study was completed using a biological replicate of Ssu72 L84F (L84F#2) since L84F samples had been consumed by this time in the research.

One of the challenges with the DiFMUP varied concentration experiment was that there were two ~ 25 kD molecular weight bands that were observed by SDS-PAGE (Figure 11) when the L84F construct was analyzed. Based on work shown in Figure 13, it was suggested that these two extra bands could be the cleaved forms of GST and Ssu72. The identity of these bands, however, was not definitive and therefore two approaches were taken for conducting the kinetic phosphatase measurements. In the first experiment, the smaller bands were included in the calculation of total Ssu72 with the assumption that all bands contributed to the phosphatase activity. In the case of the second calculation, only the band at around 50 kD was included in the calculation with the assumption that only this fraction had phosphatase activity (shown in Figure 12).[b]

The first DiFMUP varied experiment used the assumption of 7.9 mg/mL for Ssu72 concentration while the second DiFMUP varied experiment assumed a concentration of 1.8 mg/ml. In the case of each experiment, the initial rate of fluorescence change was measured for each concentration of substrate. As shown in Figure 22 and Figure 23 (with concentrations in Table 12), the L84A replicates had a much higher rate of RFU change when compared to the other three constructs. While the mean response for L84F#2 was also greater than wild type in every data point, there was overlap between points when the standard deviation was accounted for.

[b] The same lane that was used in the first calculation was used in this second calculation despite the fact that the intensity of the 50 kD band was less than the bottom intensity on the BSA calibration curve. If the lane to the left would have been used, the calculated concentration would have been 2.3 mg/mL.

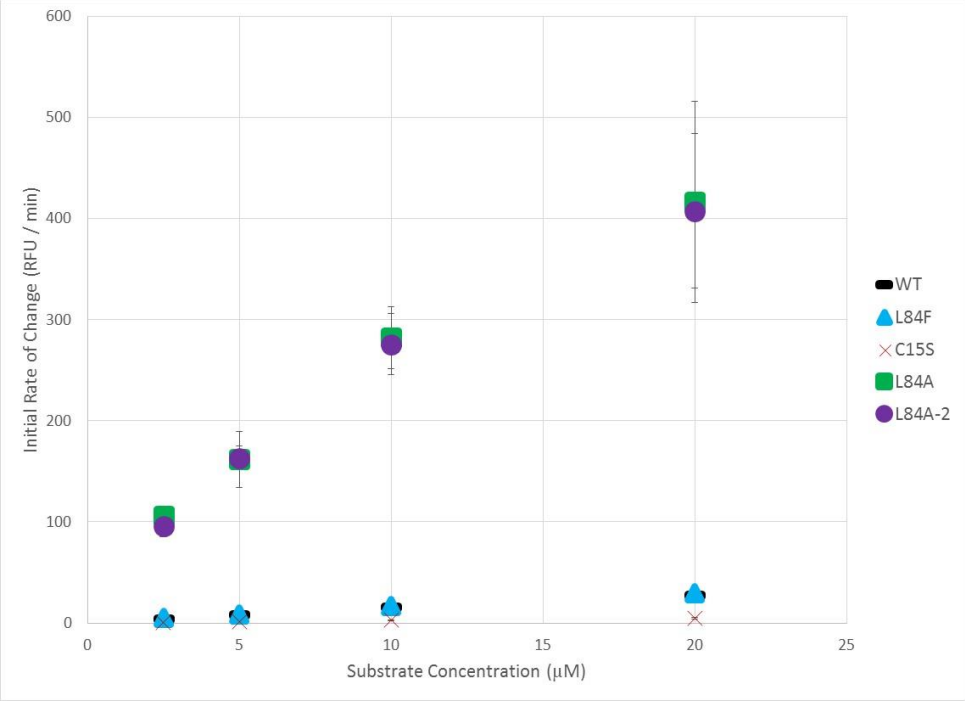


Figure 22. Kinetic study #3, L84F#2 = 7.9 mg/mL (1 standard deviation shown), with L84A

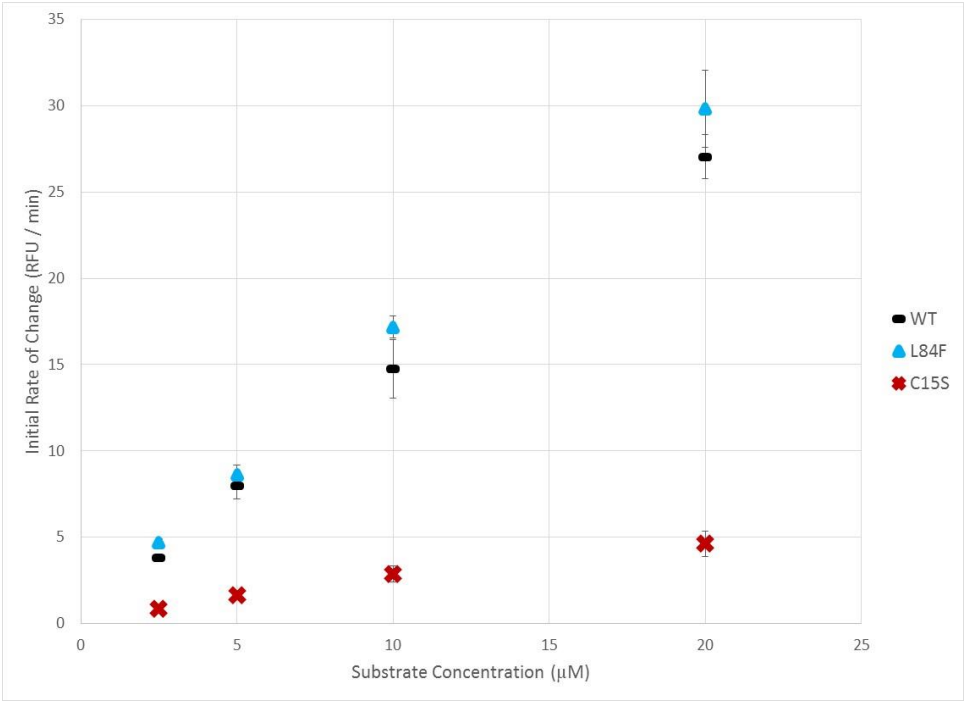


Figure 23. Kinetic study #3, L84F#2 = 7.9 mg/mL (1 standard deviation shown), without L84A

Table 12. Concentrations in Kinetic Study #3

Construct	Assumed Concentration (µg/µL)
GST-WT	13.8
GST-C15S	9.9
GST-L84A	9.1
GST-L84A2	4.9
GST-L84F2	7.9

Figure 24 below (along with corresponding concentrations in Table 13) shows the results from the analysis that assumed a lower concentration of 1.8 mg/mL. These results show clear separation and higher rate of activity for the L84F#2 construct than for wild type.

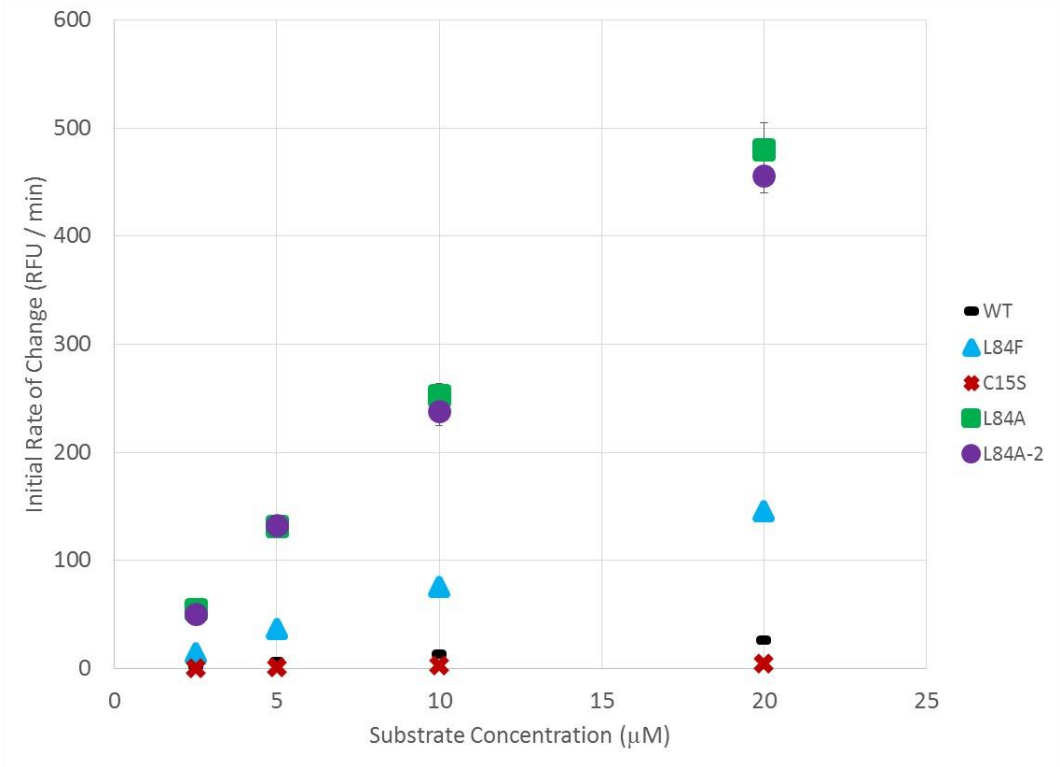


Figure 24. Kinetic Study #4, L84F#2 = 1.8 mg/mL (1 standard deviation shown)

Table 13. Concentrations in Kinetic Study #4

Construct	Assumed Concentration (µg/µL)
GST-WT	13.8
GST-C15S	9.9
GST-L84A	9.1
GST-L84A2	4.9
GST-L84F2	1.8

VI. Biotinylated Pulldown

The overall goal of the biotinylated pulldown assay was to determine whether the recombinantly produced Ssu72 constructs would associate with a CTD peptide. More specifically, there was a goal of determining whether the strength of interaction between the CTD peptide and the wild type Ssu72 differed clearly from the L84F Ssu72 construct. Results from SDS-PAGE analysis using Coomassie staining are shown in Figure 25. In the case of each supernatant lane, the bands present near 50 kD and 25 kD (in the case of the cleaved WT and GST) suggest that there were fairly consistent levels of Ssu72 protein in each sample as it was combined with the biotinylated CTD (lanes labeled as “Supernatant” or “Super”). The release step lanes (labeled as “Release”) show the levels of Ssu72 protein that were present after release of the protein from the biotinylated peptide. Each of these bands near 50 kD and 25 kD were very faint.

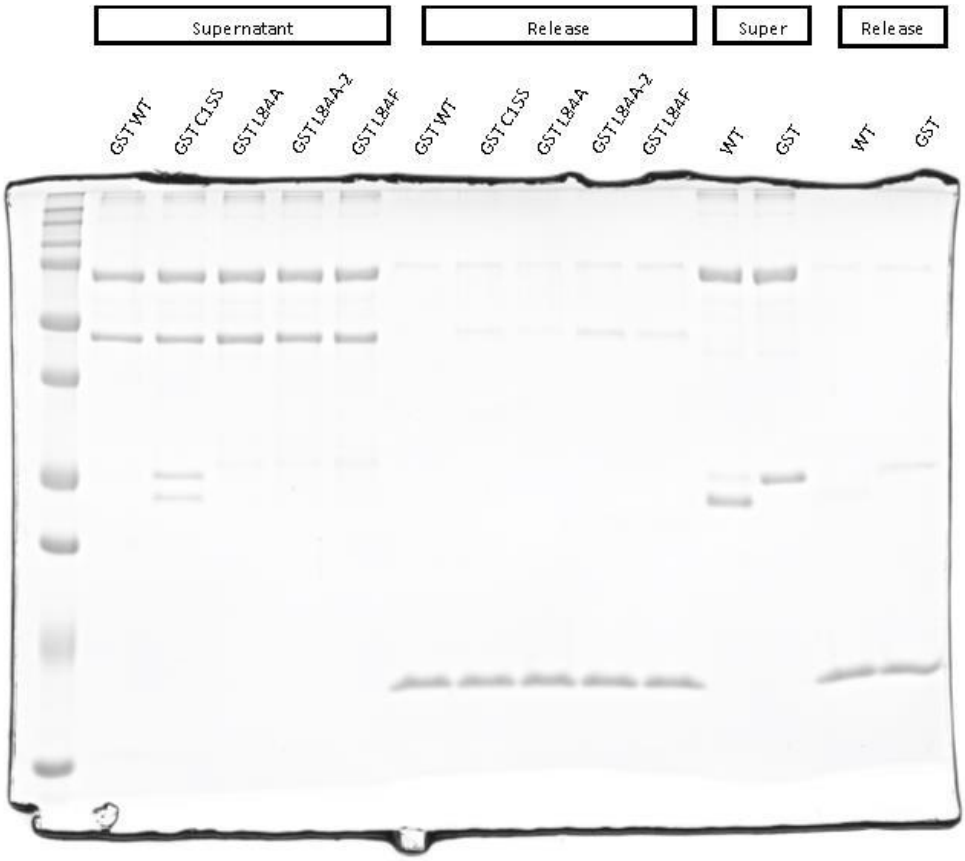


Figure 25. Biotinylated pulldown analysis by SDS PAGE and Coomassie stain

Due to the fact that the bands were difficult to see in the visualized gel, the picture file was also analyzed using the ImageJ software in order to quantitate the color density of these bands as detected by the scanner. The results from this analysis are shown in Figure 26 and indicate that in the case of every construct, there was a measurable level of association between the CTD peptide and the Ssu72 protein.

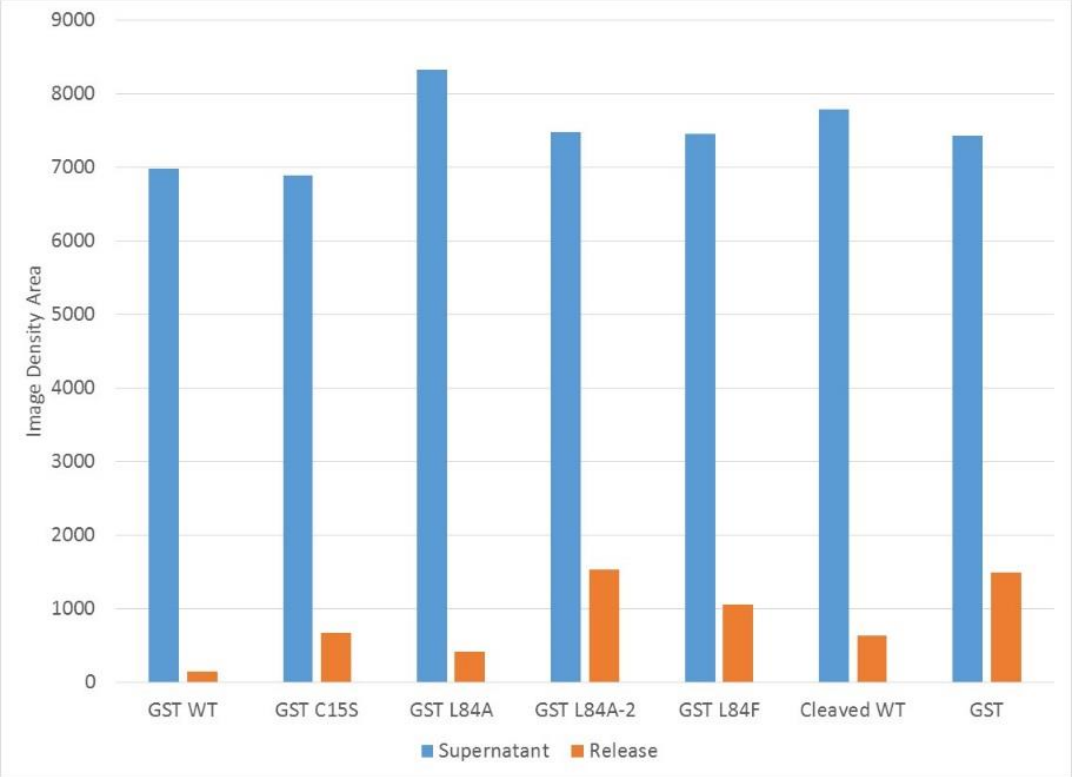


Figure 26. ImageJ Analysis of Pulldown Gel

VII. Analysis of FLAG Tagged Ssu72 Protein by SDS-PAGE and Silver Stain

The 3X FLAG Ssu72 proteins isolated from the expression and purification were first analyzed by SDS-PAGE with silver staining. The first ethanol wash consisted of 100% ethanol instead of the diluted formulation. After the ethanol wash, a water wash was done followed by the planned 30% ethanol formulation followed by the balance of the planned wash steps. Elutions were further analyzed by SDS-PAGE gel and silver staining; silver stained gels from the 3X FLAG Ssu72 Wild Type construct are shown in Figure 27 which shows a large number of bands from each of elution. Results below suggest that the expression and purification was successful from the standpoint that a large number of proteins were isolated from the procedure.

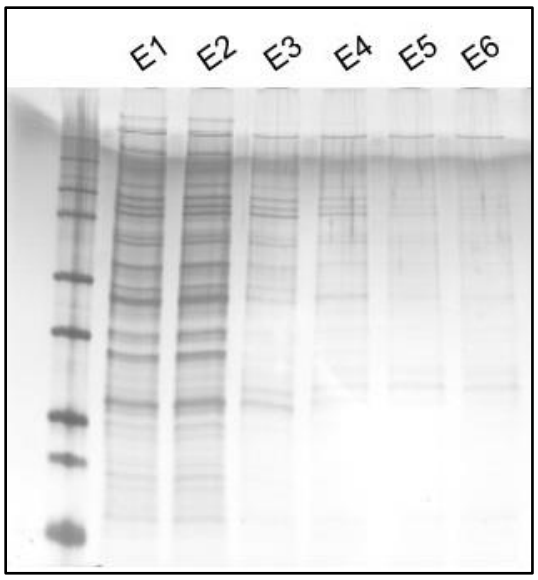


Figure 27. Elutions from 3X FLAG tagged Wild Type Ssu72

VIII. Analysis of FLAG-Ssu72 WT by LC-MS/MS

FLAG-tagged Ssu72 elutions 2 and 3 from the purification described in the section above were TCA precipitated and the precipitated protein was reduced and alkylated by TCEP and CAM. The protein was then digested by Endoproteinase Lys-C and Trypsin Gold followed by quenching with formic acid. The sample was loaded on a fused silica glass column was packed with reverse phase C18, 2.5 cm of strong cation exchange resin and 2.25 cm of C18. A10 step MudPIT method was performed on the column and data collected was analyzed by Scaffold. Steps 2 through 4 of the data set were queried further while using a 5.0% peptide false discovery rate (FDR), 95% protein threshold and 2 minimum peptides to identify a protein.

The Ssu72 protein was identified in the analysis indicating that the expression afforded the desired protein. One of the Ssu72 peptides detected with a high correlation was

VLQEAGYNVSSYGTGSAVR; the MS/MS fragment ion spectrum for this peptide is shown below in Figure 28.

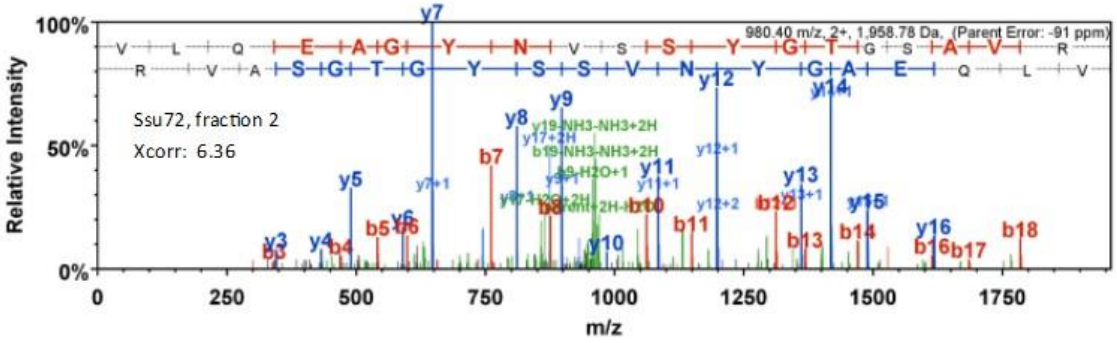


Figure 28. MS/MS spectrum for Ssu72 peptide fragment detected in fraction #2 (VLQEAGYNVSSYGTGSAVR)

Many proteins have been previously shown to be associated with the Ssu72 phosphatase *in-vivo*. Proteins that have been reported to physically associate with a TAP tagged Ssu72 are Cft1, Cft2, Ysh1, Pta1, Ref2, Pap1, Mpe1, Pfs2, Fip1, Pti1, Swd2, Glc7, Yth1 and Syc1(E. Nedea et al., 2003); each components of the cleavage and polyadenylation factor (CPF) complex. Peptides from five of these 14 proteins were detected in the LC-MS/MS analysis and these spectra along with their correlations are shown below in Figure 29 through Figure 33. Overall these results suggested that the MudPIT experiment was successful in that it detected peptides for Ssu72 along with other components of the cleavage and polyadenylation factor complex. At the same time, protein concentrations were low and would need to be increased in order to make better conclusions about what subunits of the CPF complex can be detected in the assay.

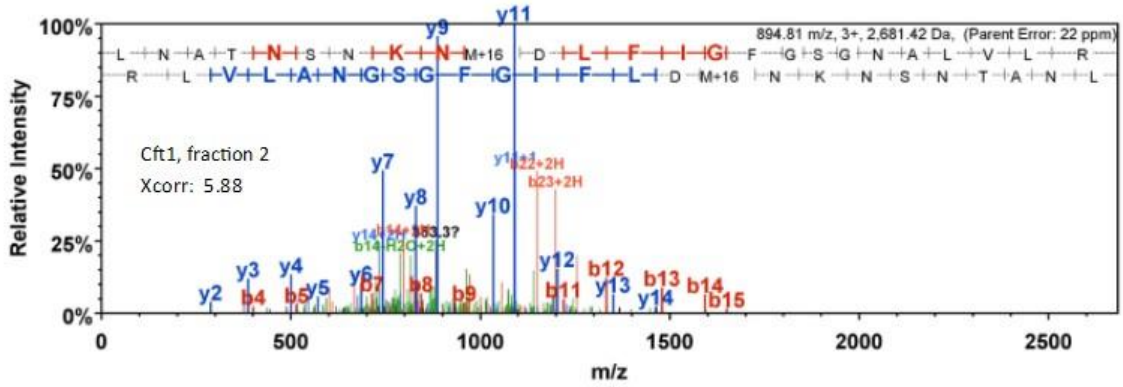


Figure 29. MS/MS spectrum for Cft1peptide detected in fraction #2

((R)LNATNSNKNmDLFIGFGSGNALVLR(L))

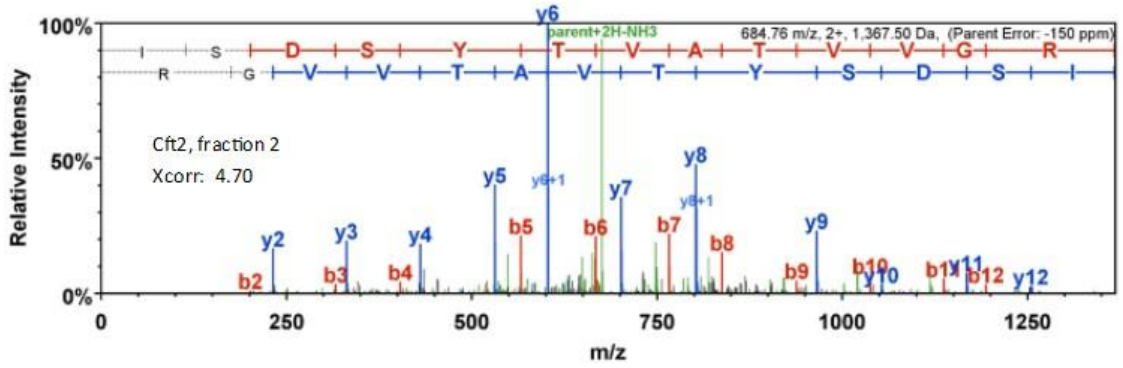


Figure 30. MS/MS spectrum for Cft2 peptide detected in fraction #2 (ISDSYTVATVVGR)

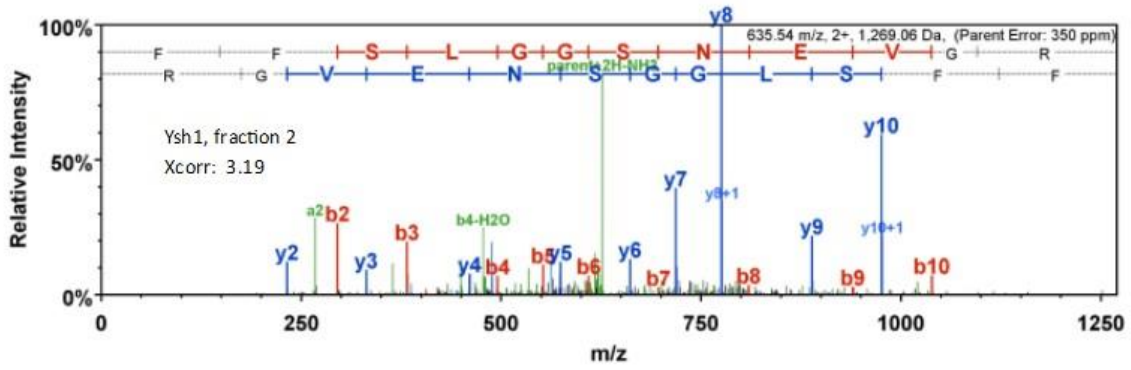


Figure 31. MS/MS spectrum for Ysh1 peptide detected in fraction #2

(FFSLGGSNEVGR)

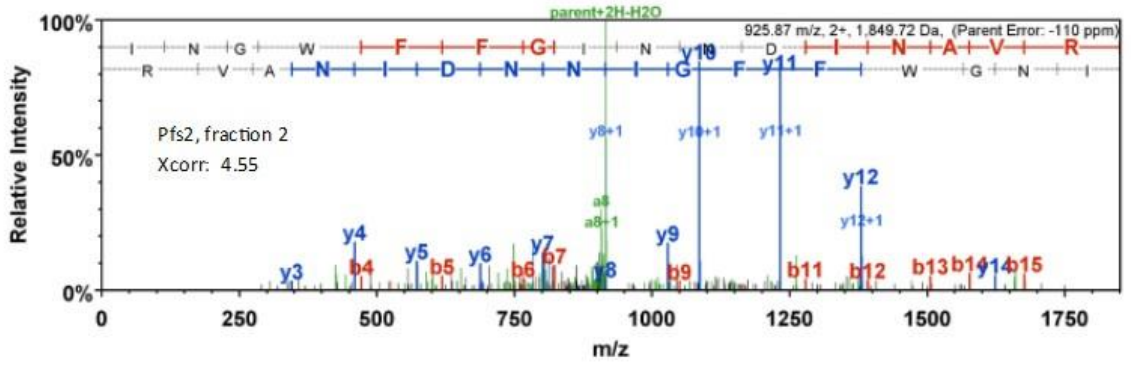


Figure 32. MS/MS spectrum for Pfs2 peptide detected in fraction #2

(INGWFFGINNDINAVR)

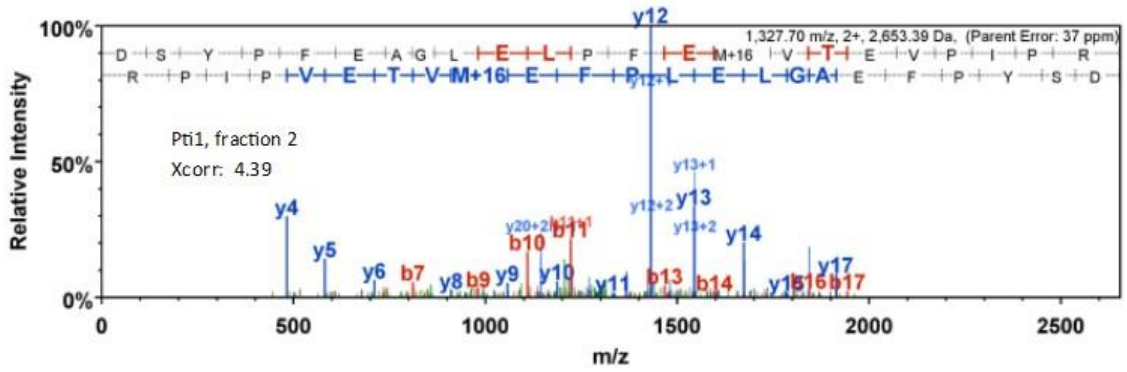


Figure 33. MS/MS spectrum for Pti1 peptide detected in fraction #2

(DSYPFEAGLELPFEMVTEVPIPR)

DISCUSSION

The overall goal of this work was to determine whether the phosphatase activity of Ssu72 TOV mutant (as measured by an *in-vitro* assay) can explain the Ssu72 L84F mutant phenotype of termination deficiency. As presented in this report, the phosphatase activity of the GST tagged Ssu72 proteins, as measured by DiFMUP substrate, is impacted by mutations made to the L84 residue. The mutation of leucine to alanine at residue 84 in general displays the highest level of phosphatase activity when compared to the other constructs tested. In experiments presented, the L84F mutation has equal or higher phosphatase activity than the wild type construct (by the *in-vitro* assay) and in no case was the measured activity lower for L84F than wild type Ssu72. The catalytically dead Ssu72 construct (cysteine to serine mutation at residue 15) shows activity lower than wild type as expected.

The increase in phosphatase activity for L84F Ssu72 is potentially the result of an improved association between Pol II CTD and Ssu72 due to pi-pi stacking interactions. This suggested relationship is based on both the observed activity by the DiFMUP substrate coupled with data from published crystal structures for the Ssu72 protein using *Drosophila melanogaster* and *Homo sapiens* homologue (Luo et al., 2013; Mayfield et al., 2015a; Werner-Allen et al., 2011; Xiang et al., 2012; Xiang et al., 2010; Y. Zhang, Zhang, & Zhang, 2011).[c]

[c] It should be noted that none of the structures published to date have used the *Saccharomyces cerevisiae* homologue due to challenges with crystallization of yeast Ssu72. In order to use the structural information published from the two available species in analysis, a sequence and structural comparison was first conducted. There is alignment between *Saccharomyces cerevisiae*, *Homo sapiens* and *Drosophila melanogaster* forms; 35% identity, 72 similar positions and 66 similar positions based on the Clustal algorithm (Soding, 2005). The alignment analysis suggested that yeast Ssu72 could be understood at a rudimentary level by the *Drosophila* and Human homologues.

Ssu72 structures published with a CTD peptide bound to the catalytic groove of the phosphatase were viewed for both the *Drosophila* and human homologues. A *Drosophila* structure (4YGX) (Mayfield et al., 2015a) and human structure (3O2Q) (Xiang et al., 2010) viewed in NCBI Cn3D 4.3.1 are shown in Figure 34.

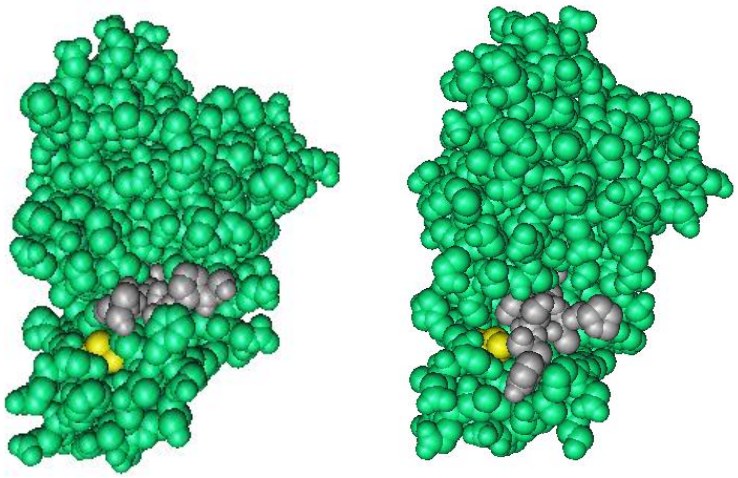


Figure 34. 4YGX (Left) and 3O2Q (Right) with Ssu72 L82 and I81 highlighted yellow and CTD peptide colored gray

The two crystal structures, 4YGX and 3O2Q, were rendered and compared using WinCoot version 0.8.2 (as shown below in Figure 35 and Figure 36) to visualize some of the interactions at the corresponding residues. As shown in these figures below, the wild type leucine or isoleucine in Ssu72 for the two species are more distant from the CTD peptide (4.4 and 4.6 Angstroms respectively) than the mutated Ssu72 phenylalanine (4.0 and 3.2 Angstroms respectively) that was generated using the mutation function in WinCoot. It is possible that the proline in the CTD is interacting with the phenylalanine by pi-pi interactions. Reports suggest that proline can participate in pi-pi stacking with other aromatic side chains (Zondlo, 2013) and that 3.4 Angstroms is a good approximation for distance of the interactions (Hunter & Sanders, 1990).

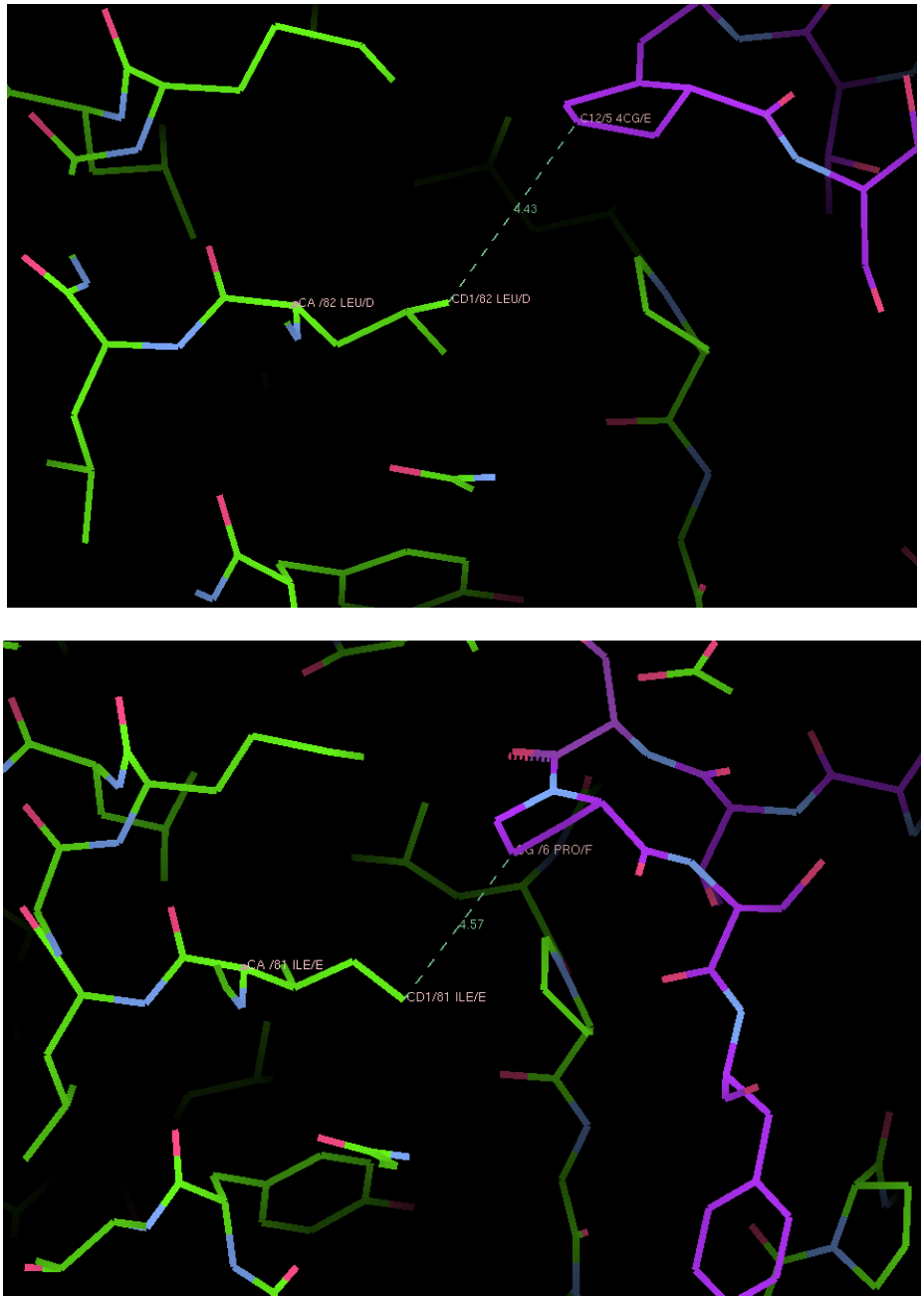


Figure 35. 4YGX, *Drosophila melanogaster* L82 residue (Top) and 3O2Q, *Homo sapiens* I81 residue (Bottom) and distance to CTD peptide in crystal structure

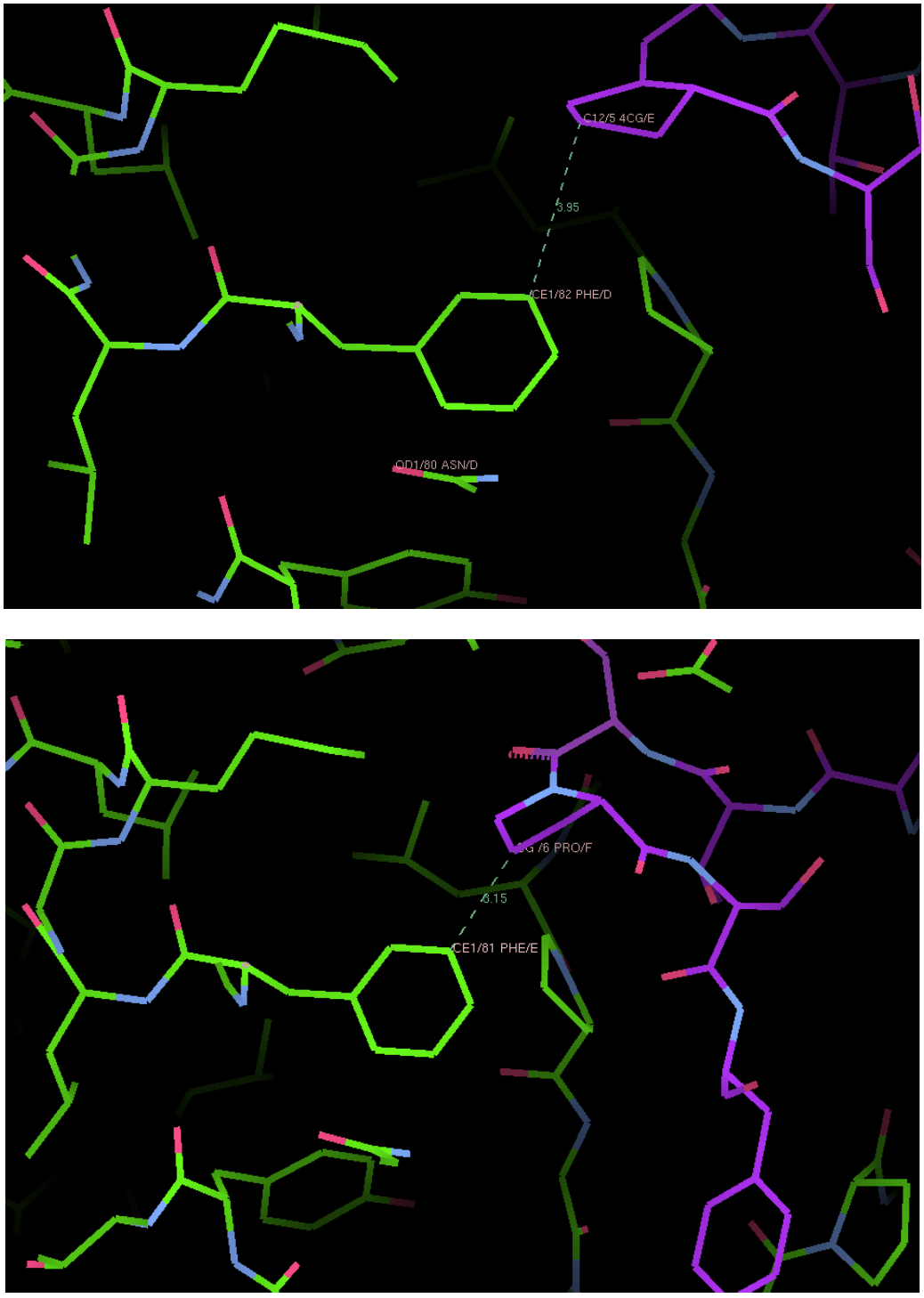


Figure 36. 4YGX, *Drosophila melanogaster* L82F mutation (Top) and 3O2Q, *Homo sapiens* I81F mutation (Bottom) and distance to CTD peptide in crystal structure

In the case of the Ssu72 L84A mutation (as shown in Figure 37), it is unclear why this mutation results in such an extremely high level of activity as measured by the DiFMUP assay. As shown below, the mutation potentially results in a greater distance between the Ssu72 residue and the CTD peptide. It is possible that the mutation affects the tertiary structure of the Ssu72 protein and causes the CTD to be more accessible to the catalytic cysteine site. On the other hand, it is possible that the L84A activity by the DiFMUP assay is not representative of the *in-vivo* response. In other words, it is quite possible that termination override would not be observed in this L84A condition. Despite this contradiction between phosphatase response and suggested interactions by crystal structures, it can be seen that the L84 residue has a good potential for impacting the interaction with the Pol II CTD.

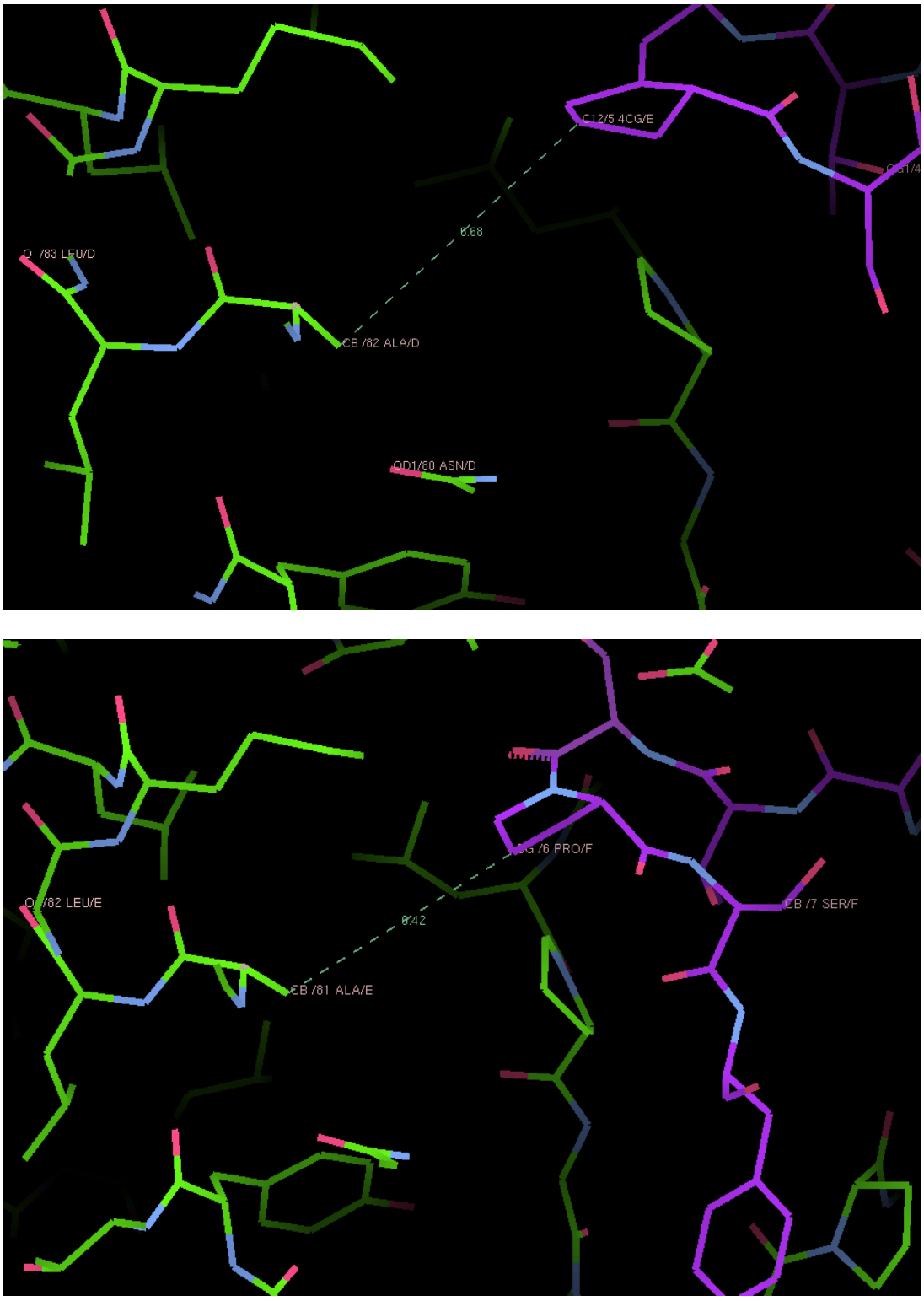


Figure 37. 4YGX, *Drosophila melanogaster* L82A mutation (Top) and 3O2Q, *Homo sapiens* I81A mutation (Bottom) and distance to CTD peptide in crystal structure

It should be noted that data from Zhang showed that L82A mutation in a *Drosophila* homologue had no loss in phosphatase activity (by PNPP mimetic substrate) or loss of CTD substrate specificity (Y. Zhang et al., 2011) when compared to wild type. The studies presented in this report suggest a possible increase in activity for L84F Ssu72 when compared to Ssu72 WT. It has also been observed in unpublished studies within the Mosley lab that serine 5 CTD phosphorylation levels were lower for a SSU72-TOV containing *S. cerevisiae* strain. On the other hand, if the L84F mutation doesn't actually increase phosphatase activity but instead maintains the same phosphatase activity as WT, it is possible that the mutation impacts some other interaction with the CPF complex or RNA Pol II. Data from the pull-down approach suggests that all of GST-Ssu72 homologues tested, including the L84F mutant, are able to associate with the CTD (based on results with a CTD peptide). At the same time, GST alone associated with the CTD peptide and reinforces the need for re-execution of the method with GST cleaved forms of Ssu72.

Another possibility is that the Ssu72 L84F mutation actually affects interactions with other CPF subunits. The results from the *ex-vivo* approach with mass spec analysis suggest that the use of a FLAG tagged Ssu72 is suitable approach for determining physical interactions between the Ssu72 protein and other CPF subunits. The presence of the five detected subunits agrees with external reports showing that Ssu72 is stably associated with the CPF complex. At the same time, the number of proteins that were observed along with their intensity indicate that the experiment could be improved with higher concentrations of protein. The utility of FLAG tagged Ssu72 expression and MudPIT analysis could further be leveraged in the future by comparing results with those from a similar experiment where the TOV mutant is used instead of the WT expressing

cell line. The mass spec comparative analysis would allow for understanding the influence of the L84F mutation on the physical interaction with other CPF components and consequently help to understand if the mutation causes other changes in the transcription process.

The ultimate goal of our work, however, is to explain the impact of the L84F mutation on the observed termination override (TOV) described by the Reines lab. One possible explanation for the observed TOV is that the mutation affects association with termination factors. Ssu72 is associated with termination of snoRNA, (M. Kim et al., 2006; Steinmetz & Brow, 2003) and the termination of these non-coding transcripts has been reported to be associated with the NNS pathway (Nrd1, Nab3, Sen1). The NNS pathway is described pictorially below in Figure 38.

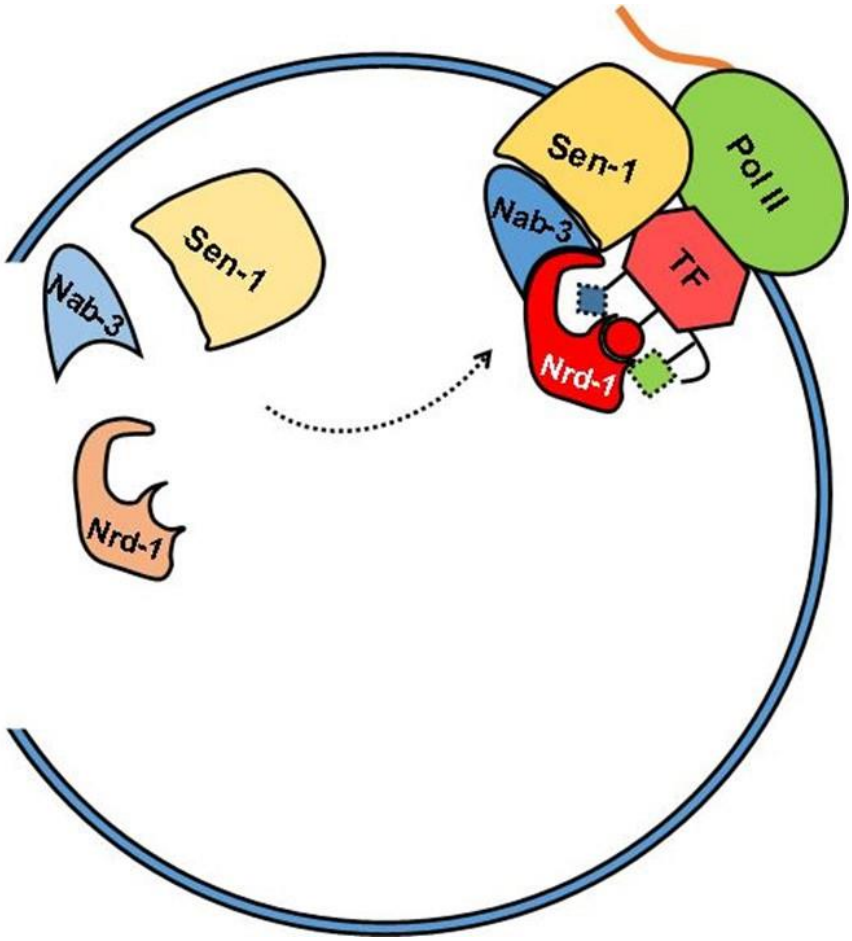


Figure 38. NNS pathway and association with DNA transcription termination by RNA Pol II

It has been reported that Nrd1 binds to Ser-5 phosphorylated CTD (Vasiljeva et al., 2008). If the phosphatase activity of the L84F mutant is slightly elevated over the wild type Ssu72, it is possible that the lower level of Ser-5 phosphorylation of the CTD results in less efficient docking of the NNS complex to the CTD. Concomitantly, if the CTD binding domain of Nrd1 cannot bind to RNA Pol II CTD, it is possible that decreased binding could lead to ineffective termination; this theory is reinforced by the fact that mutations of Nrd1 have been connected with termination read-through in other reports (Vasiljeva & Buratowski, 2006). Read-through was also reported when the C-terminal interacting domain (CID) of Nrd1 was mutated (Heo et al., 2013). Overall, if it is assumed

that the activity of L84F increases the phosphatase activity of Ssu72 when compared to wild type Ssu72, the model for termination override can be described by Figure 39 below.

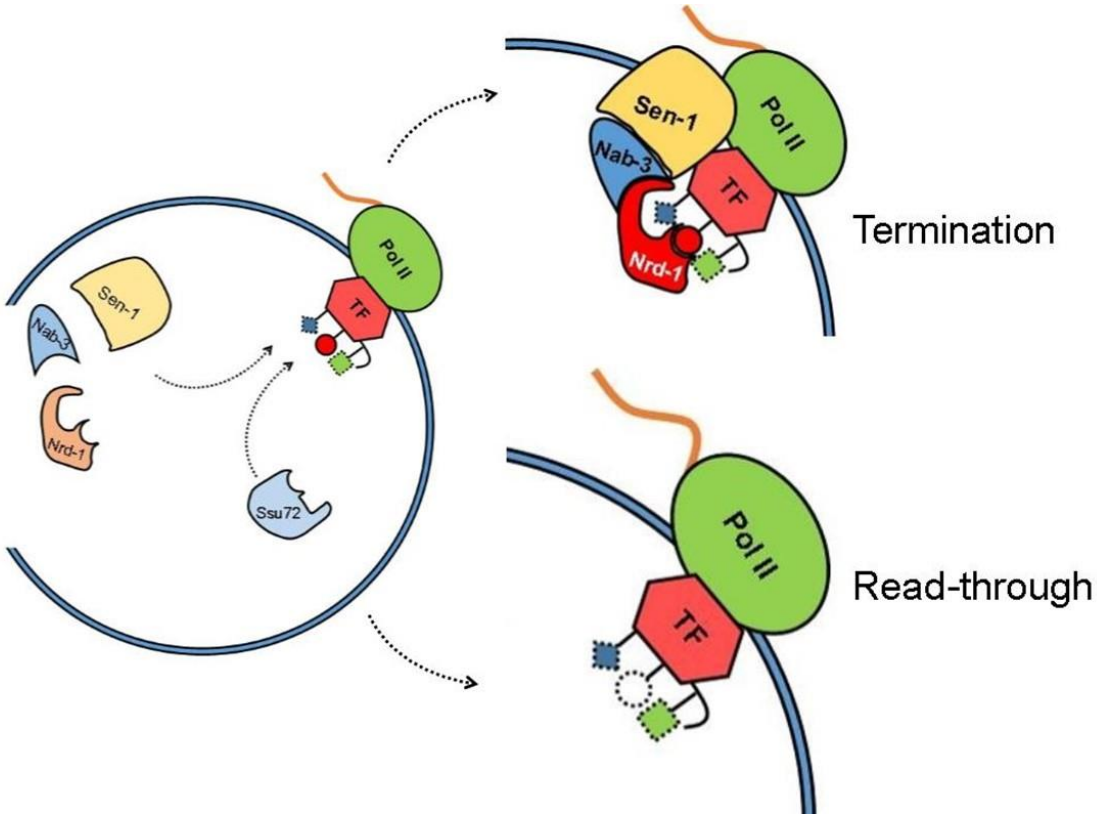


Figure 39. Model of transcription termination and read-through via interactions with both NNS and CPF/Ssu72

CONCLUSIONS

It can be concluded from this work that the L84F mutation in the Ssu72 yeast isoform causes an equivalent or increase in *in-vitro* phosphatase activity as compared to wild type. It was also shown from this work that the L84A mutation in the yeast isoform causes an even greater increase in *in-vitro* phosphatase activity and further reinforces that the L84 residue influences activity of the phosphatase. One theory for the increased activity observation is that the leucine to phenylalanine mutation improves Ssu72 substrate access. This may be true for both small molecule substrates such as DiFMUP as well as the RNA Pol II CTD. The ultimate goal of this work, however, is to explain the impact of the L84F mutation on the observed termination override (TOV) mutant of Ssu72. One possible explanation for the observed TOV is that the L84F mutation results in lower serine 5 phosphorylation (due to higher phosphatase activity when compared to wild type Ssu72) which decreases the association of termination factors such as Nrd1 with Pol II (leading to inefficient termination). The conclusions highlighted in this report could be improved with further suggested work. It is recommended that more proteomics work be used to orthogonally describe the TOV mutant, which would allow for understanding whether the L84F mutation impacts the association of the phosphatase with other CPF subunits. It would also be beneficial to crystallize the Ssu72 mutant constructs in the presence of a CTD peptide to confirm whether the observed phosphatase activity can be supported. Human and *Drosophila* homologues have already been crystallized and it would be advantageous to use this existing work (since molecular replacement and existing crystallization conditions could be leveraged).

REFERENCES

Archambault, J., Chambers, R. S., Kobor, M. S., Ho, Y., Cartier, M., Bolotin, D., . . . Greenblatt, J. (1997). An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences, 94(26), 14300-14305.

Barilla, D., Lee, B. A., & Proudfoot, N. J. (2001). Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America, 98(2), 445-450. doi:10.1073/pnas.021545298

Bataille, A. R., Jeronimo, C., Jacques, P. E., Laramee, L., Fortin, M. E., Forest, A., . . . Robert, F. (2012). A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Mol Cell, 45(2), 158-170. doi:10.1016/j.molcel.2011.11.024

Bentley, D. (2002). The mRNA assembly line: transcription and processing machines in the same factory. Curr Opin Cell Biol, 14(3), 336-342.

Breathnach, R., & Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem, 50, 349-383. doi:10.1146/annurev.bi.50.070181.002025

Brown, S. A., Imbalzano, A. N., & Kingston, R. E. (1996). Activator-dependent regulation of transcriptional pausing on nucleosomal templates. Genes Dev, 10(12), 1479-1490.

Buratowski, S. (2003). The CTD code. Nat Struct Mol Biol, 10(9), 679-680.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol Cell, 36(4), 541-546. doi:10.1016/j.molcel.2009.10.019

Chapman, R. D., Heidemann, M., Albert, T. K., Mailhammer, R., Flatley, A., Meisterernst, M., . . . Eick, D. (2007). Transcribing RNA Polymerase II Is Phosphorylated at CTD Residue Serine-7. Science, 318(5857), 1780-1782. doi:10.1126/science.1145977

Chen, F., Zhou, Y., Qi, Y. B., Khivansara, V., Li, H., Chun, S. Y., . . . Jin, Y. (2015). Context-dependent modulation of Pol II CTD phosphatase SSUP-72 regulates alternative polyadenylation in neuronal development. Genes Dev, 29(22), 2377-2390. doi:10.1101/gad.266650.115

Chen, Y., Zhang, L., Estarás, C., Choi, S. H., Moreno, L., Karn, J., . . . Jones, K. A. (2014). A gene-specific role for the Ssu72 RNAPII CTD phosphatase in HIV-1 Tat transactivation. Genes Dev, 28(20), 2261-2275. doi:10.1101/gad.250449.114

Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., & Dahmus, M. E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proceedings of the National Academy of Sciences of the United States of America, 82(23), 7934-7938.

Cramer, P., Armache, K. J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., . . . Vannini, A. (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys, 37, 337-352. doi:10.1146/annurev.biophys.37.032807.130008

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., & Keller, W. (2002). A Role for SSU72 in Balancing RNA Polymerase II Transcription Elongation and Termination. Molecular Cell, 10(5), 1139-1150.

Dichtl, B., Blank, D., Sadowski, M., Hubner, W., Weiser, S., & Keller, W. (2002). Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. Embo j, 21(15), 4125-4135.
Florens, L., & Washburn, M. P. (2006). Proteomic analysis by multidimensional protein identification technology. Methods Mol Biol, 328, 159-175. doi:10.1385/1-59745-026-x:159
Heo, D.-h., Yoo, I., Kong, J., Lidschreiber, M., Mayer, A., Choi, B.-Y., . . . Kim, M. (2013). The RNA Polymerase II C-terminal Domain-interacting Domain of Yeast Nrd1 Contributes to the Choice of Termination Pathway and Couples to RNA Processing by the Nuclear Exosome. Journal of Biological Chemistry, 288(51), 36676-36690. doi:10.1074/jbc.M113.508267
Hsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev, 26(19), 2119-2137. doi:10.1101/gad.200303.112
Hunter, C. A., & Sanders, J. K. M. (1990). The nature of .pi.-.pi. interactions. Journal of the American Chemical Society, 112(14), 5525-5534. doi:10.1021/ja00170a016
Jeronimo, C., Bataille, A. R., & Robert, F. (2013). The writers, readers, and functions of the RNA polymerase II C-terminal domain code. Chem Rev, 113(11), 8491-8522. doi:10.1021/cr4001397
Kim, M., Vasiljeva, L., Rando, O. J., Zhelkovsky, A., Moore, C., & Buratowski, S. (2006). Distinct pathways for snoRNA and mRNA termination. Mol Cell, 24(5), 723-734. doi:10.1016/j.molcel.2006.11.011
Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., . . . Ren, B. (2005). A high-resolution map of active promoters in the human genome. Nature, 436(7052), 876-880. doi:10.1038/nature03877
Kim, T. K., Ebright, R. H., & Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science, 288(5470), 1418-1422.
Komarnitsky, P., Cho, E.-J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev, 14(19), 2452-2460.
Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev, 14(19), 2452-2460.
Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., & Hampsey, M. (2004). Ssu72 Is an RNA Polymerase II CTD Phosphatase. Molecular Cell, 14(3), 387-394. doi:10.1016/s1097-2765(04)00235-7
Kyburz, A., Sadowski, M., Dichtl, B., & Keller, W. (2003). The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. Nucleic Acids Res, 31(14), 3936-3945.
Licatalosi, D. D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J. B., & Bentley, D. L. (2002). Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. Mol Cell, 9(5), 1101-1111.
Liu, H., Dibling, B., Spike, B., Dirlam, A., & Macleod, K. (2004). New roles for the RB tumor suppressor protein. Current Opinion in Genetics & Development, 14(1), 55-64.
Loya, T. J., O'Rourke, T. W., & Reines, D. (2012). A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3. Nucleic Acids Res, 40(15), 7476-7491. doi:10.1093/nar/gks377
Luo, Y., Yogesha, S. D., Cannon, J. R., Yan, W., Ellington, A. D., Brodbelt, J. S., & Zhang, Y. (2013). novel modifications on C-terminal domain of RNA polymerase

II can fine-tune the phosphatase activity of Ssu72. ACS Chem Biol, 8(9), 2042-2052. doi:10.1021/cb400229c
Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J., & Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol, 17(10), 1272-1278.
Mayfield, J. E., Fan, S., Wei, S., Zhang, M., Li, B., Ellington, A. D., . . . Zhang, Y. (2015a). Chemical tools to decipher regulation of phosphatases by proline isomerization on eukaryotic RNA polymerase II. ACS Chem Biol. doi:10.1021/acschembio.5b00296
Mayfield, J. E., Fan, S., Wei, S., Zhang, M., Li, B., Ellington, A. D., . . . Zhang, Y. J. (2015b). Chemical Tools To Decipher Regulation of Phosphatases by Proline Isomerization on Eukaryotic RNA Polymerase II. ACS Chem Biol, 10(10), 2405-2414. doi:10.1021/acschembio.5b00296
McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., . . . Bentley, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev, 11(24), 3306-3318.
McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., . . . Bentley, D. L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature, 385(6614), 357-361. doi:10.1038/385357a0
Meinhart, A., Silberzahn, T., & Cramer, P. (2003). The mRNA transcription/processing factor Ssu72 is a potential tyrosine phosphatase. J Biol Chem, 278(18), 15917-15921. doi:10.1074/jbc.M301643200
Morgan, G. T. (2002). Lampbrush chromosomes and associated bodies: new insights into principles of nuclear structure and function. Chromosome Res, 10(3), 177-200.
Mosley, A. L., Pattenden, S. G., Carey, M., Venkatesh, S., Gilmore, J. M., Florens, L., . . . Washburn, M. P. (2009). Rtr1 is a CTD Phosphatase that Regulates the Transition from Serine 5 to Serine 2 Phosphorylation during the RNA Polymerase II Transcription Cycle. Molecular Cell, 34(2), 168-178. doi:10.1016/j.molcel.2009.02.025
Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., . . . Greenblatt, J. (2003). Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. J Biol Chem, 278(35), 33000-33010. doi:10.1074/jbc.M304454200
Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., . . . Greenblatt, J. (2003). Organization and Function of APT, a Subcomplex of the Yeast Cleavage and Polyadenylation Factor Involved in the Formation of mRNA and Small Nucleolar RNA 3′-Ends. Journal of Biological Chemistry, 278(35), 33000-33010. doi:10.1074/jbc.M304454200
O'Shea-Greenfield, A., & Smale, S. T. (1992). Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. J Biol Chem, 267(2), 1391-1402.
Pappas, D. L., Jr., & Hampsey, M. (2000). Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in Saccharomyces cerevisiae. Mol Cell Biol, 20(22), 8343-8351.
Phatnani, H. P., & Greenleaf, A. L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev, 20(21), 2922-2936. doi:10.1101/gad.1477006

Porrua, O., & Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nat Rev Mol Cell Biol, 16(3), 190-202. doi:10.1038/nrm3943

Richard, P., & Manley, J. L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev, 23(11), 1247-1269. doi:10.1101/gad.1792809

Riethoven, J. J. (2010). Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. Methods Mol Biol, 674, 33-42. doi:10.1007/978-1-60761-854-6_3

Roeder, R. G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends in Biochemical Sciences, 16, 402-408.

Saunders, A., Core, L. J., & Lis, J. T. (2006). Breaking barriers to transcription elongation. Nat Rev Mol Cell Biol, 7(8), 557-567. doi:10.1038/nrm1981

Sims, R. J., 3rd, Belotserkovskaya, R., & Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev, 18(20), 2437-2468. doi:10.1101/gad.1235904

Smale, S. T., Schmidt, M. C., Berk, A. J., & Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proceedings of the National Academy of Sciences of the United States of America, 87(12), 4509-4513.

Smith-Kinnaman, W. R., Berna, M. J., Hunter, G. O., True, J. D., Hsu, P., Cabello, G. I., . . . Mosley, A. L. (2014). The interactome of the atypical phosphatase Rtr1 in Saccharomyces cerevisiae. Molecular bioSystems, 10(7), 1730-1741. doi:10.1039/c4mb00109e

Soding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics, 21(7), 951-960. doi:10.1093/bioinformatics/bti125

St-Pierre, B., Liu, X., Kha, L. C., Zhu, X., Ryan, O., Jiang, Z., & Zacksenhaus, E. (2005). Conserved and specific functions of mammalian ssu72. Nucleic Acids Res, 33(2), 464-477. doi:10.1093/nar/gki171

Steinmetz, E. J., & Brow, D. A. (2003). Ssu72 Protein Mediates Both Poly(A)-Coupled and Poly(A)-Independent Termination of RNA Polymerase II Transcription. Mol Cell Biol, 23(18), 6339-6349. doi:10.1128/MCB.23.18.6339-6349.2003

Sun, Z. W., & Hampsey, M. (1996). Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. Mol Cell Biol, 16(4), 1557-1566.

Tan-Wong, S. M., Zaugg, J. B., Camblong, J., Xu, Z., Zhang, D. W., Mischo, H. E., . . . Proudfoot, N. J. (2012). Gene loops enhance transcriptional directionality. Science, 338(6107), 671-675. doi:10.1126/science.1224350

Vasiljeva, L., & Buratowski, S. (2006). Nrd1 Interacts with the Nuclear Exosome for 3′ Processing of RNA Polymerase II Transcripts. Molecular Cell, 21(2), 239-248.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., & Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol, 15(8), 795-804.

Wang, W., Carey, M., & Gralla, J. D. (1992). Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. Science, 255(5043), 450-453.

Werner-Allen, J. W., Lee, C. J., Liu, P., Nicely, N. I., Wang, S., Greenleaf, A. L., & Zhou, P. (2011). cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. J Biol Chem, 286(7), 5717-5726. doi:10.1074/jbc.M110.197129

Xiang, K., Manley, J. L., & Tong, L. (2012). An unexpected binding mode for a Pol II CTD peptide phosphorylated at Ser7 in the active site of the CTD phosphatase Ssu72. Genes Dev, 26(20), 2265-2270. doi:10.1101/gad.198853.112

Xiang, K., Nagaike, T., Xiang, S., Kilic, T., Beh, M. M., Manley, J. L., & Tong, L. (2010). Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. Nature, 467(7316), 729-733. doi:10.1038/nature09391

Yeo, M., Lin, P. S., Dahmus, M. E., & Gill, G. N. (2003). A Novel RNA Polymerase II C-terminal Domain Phosphatase That Preferentially Dephosphorylates Serine 5. Journal of Biological Chemistry, 278(28), 26078-26085. doi:10.1074/jbc.M301791200

Yudkovsky, N., Ranish, J. A., & Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. Nature, 408(6809), 225-229. doi:10.1038/35041603

Zawel, L., Kumar, K. P., & Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. Genes Dev, 9(12), 1479-1490.

Zhang, D. W., Mosley, A. L., Ramisetty, S. R., Rodríguez-Molina, J. B., Washburn, M. P., & Ansari, A. Z. (2012). Ssu72 Phosphatase-dependent Erasure of Phospho-Ser7 Marks on the RNA Polymerase II C-terminal Domain Is Essential for Viability and Transcription Termination. J Biol Chem, 287(11), 8541-8551. doi:10.1074/jbc.M111.335687

Zhang, Y., Zhang, M., & Zhang, Y. (2011). Crystal structure of Ssu72, an essential eukaryotic phosphatase specific for the C-terminal domain of RNA polymerase II, in complex with a transition state analogue. Biochem J, 434(3), 435-444. doi:10.1042/bj20101471

Zondlo, N. J. (2013). Aromatic-Proline Interactions: Electronically Tunable CH/π Interactions. Accounts of chemical research, 46(4), 1039-1049. doi:10.1021/ar300087y

CURRICULUM VITAE

Neil Andrew McCracken

Education: **B.S., Chemical Engineering,** Rose-Hulman Institute of Technology, 2000.

Licenses: Engineering – State of Indiana

Publications: McCracken, N. A., et al. (2014). "Control of galactosylated glycoforms distribution in cell culture system." Biotechnol Prog **30**(3): 547-553.

Professional Experience:

ELI LILLY AND COMPANY **August 2008 - Present**
CORPORATE TECHNOLOGY CENTER
Bioproduct R&D (November 2010 – Present)

- Responsible for development of novel mammalian cell culture processes.
  - Advanced department understanding of cell culture variables that affect antibody post- translational modifications such as glycosylation.
  - Successfully developed processes for several bioproduct candidates that achieved first human dose.
  - Developed a process using mathematical modeling of cell growth secretion and cell protease clipping of the secreted therapeutic.
- Promotion History: Promoted to Consultant Engineer 2013

Chemical Product R&D (August 2008 – November 2010)

- Responsible for development and technology transfer of GMP processes for novel late phase small molecule intermediates and API.
  - Utilized reaction engineering principles to create mathematical models that describes desired and undesired reactions for a process thus allowing for the understanding of the process design space and modeling of potential process improvements.
  - Completed technology transfer of a process to a future manufacturing site (located overseas).

W.L. GORE & ASSOCIATES **August 2006 - June 2007**
MEDICAL PRODUCTS DIVISION
Vascular Graft Production (August 2006 – June 2007)

- Responsible for support of vascular graft product line along with new product lines.
  - Analyzed and addressed seven potential non-conforming product cases.

- Successfully implemented three changes that streamlined existing processes.
- Used existing data to successfully implement two final product testing improvements.
- Validated equipment for new vascular graft related product line.

ELI LILLY AND COMPANY **June 2000 - July 2006**
CORPORATE TECHNOLOGY CENTER
Chemical Product R&D (August 2003 – July 2006)

- Responsible for laboratory development, pilot plant scale-up and analysis of GMP processes for novel early and late phase small molecule intermediates and API.
- Responsible for the oversight and technical development of an engineering technician.
- Promotion History: Promoted to Process Engineer in 2003 and Senior Process Engineer in 2006

ELI LILLY AND COMPANY – CLINTON LABORATORIES
Environmental Controls (June 2000 – August 2003)

- Responsible for troubleshooting, compliance and upgrade of five hazardous waste thermal oxidizers.
- Responsible for development and implementation of new business processes and procedures within the department designed to ensure safety and compliance.

THE SHEPHERD COLOR COMPANY **1998, 1999, 2000**
Summer Laboratory Co-op

- Analyzed pigment shipped globally for compliance with North American/European standards utilizing DCP, atomic absorption, surface area, acid/base extraction, color metric, pH methods.
- Tested waste water produced by plant for heavy metal/ammonia content.